

STUDY OF DIFFERENT TYPES OF GAUSSIAN FILTERS
(KALMAN ,EXTENDED KALMAN ,UNSCENTED,EXTENDED
COMPLEX KALMAN FILTERS)

*A project report submitted in partial fulfillment of the requirements for
the degree of*

Bachelor of Technology in Electrical Engineering

ABHISEK MOHANTY

SIDHARTHA KAR



National Institute of Technology, Rourkela

Rourkela-769008, Orissa.



National Institute of Technology, Rourkela

Rourkela-769008, Orissa.

ACKNOWLEDGEMENT

We would like to express our gratitude towards all the people who have contributed their precious time and effort to help us. Without them it would not have been possible for us to understand and complete the project.

We would like to thank **Professor S.RAUTA Department of Electrical Engineering**, our Project Supervisor for his guidance, support, motivation and encouragement through out the period this work was carried out. His readiness for consultation at all times, his educative comments, his concern and assistance even with practical things have been invaluable.

We would also like to express our heart-felt gratitude to the faculty of Electrical Engineering Department for providing a challenging and enlightening experience. We remain indebted to National Institute of Technology, Rourkela for providing us this platform of excellence.

A project of this nature could never have been attempted without our reference to and inspiration from the works of others whose details are mentioned in the References section. We acknowledge our indebtedness to all of them. Last, but not the least, our sincere thanks to all of our friends who have patiently extended all sorts of help for accomplishing this undertaking

Date:

ABHISEK MOHANTY(10502006)
SIDHARTHA KAR (10502007)



**Department of Electrical Engineering,
National Institute of Technology, Rourkela.**

Rourkela-769008, Orissa.

CERTIFICATE

This is to certify that the thesis entitled, “**STUDY OF DIFFERENT TYPES OF GAUSSIAN FILTERS(KALMAN ,EXTENDED KALMAN ,UNSCENTED AND EXTENDED COMPLEX KALMAN FILTERS)**” submitted by **ABHISEK MOHANTY** for the partial fulfillment of the requirements for the award of Bachelor of Technology degree in Electrical Engineering at the National Institute of Technology, Rourkela (Deemed University) is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge the matter embodied in the thesis has not been submitted to any other University/Institute for the award of any degree or diploma.

Date:

professor S .RAUTA

Department of Electrical Engineering,

National Institute of Technology, Rourkela

Rourkela-769008, Orissa.



**Department of Electrical Engineering,
National Institute of Technology, Rourkela.**

Rourkela-769008, Orissa.

CERTIFICATE

This is to certify that the thesis entitled, “**STUDY OF DIFFERENT TYPES OF GAUSSIAN FILTERS(KALMAN ,EXTENDED KALMAN ,UNSCENTED AND EXTENDED COMPLEX KALMAN FILTERS)**” submitted by **SIDHARTHA KAR** for the partial fulfillment of the requirements for the award of Bachelor of Technology degree in Electrical Engineering at the National Institute of Technology, Rourkela (Deemed University) is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge the matter embodied in the thesis has not been submitted to any other University/Institute for the award of any degree or diploma.

Date:

Professor S .RAUTA

Department of Electrical Engineering,

National Institute of Technology, Rourkela

Rourkela-769008, Orissa.

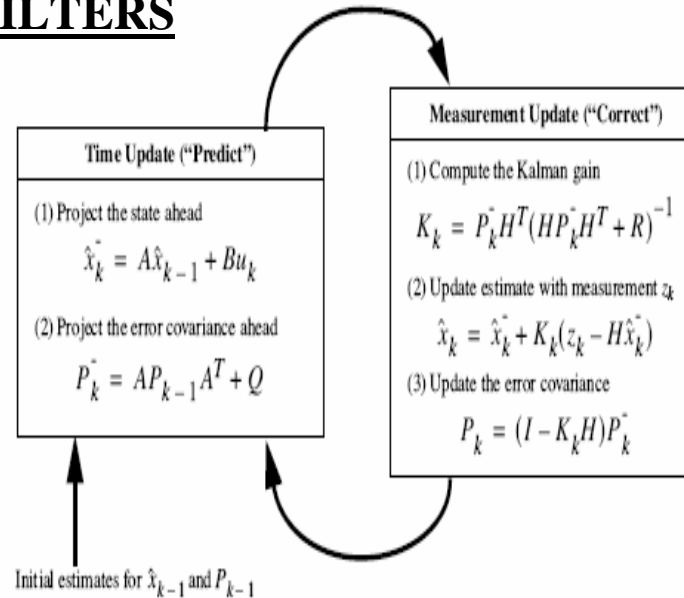
ABSTRACT

Filtering and estimation are two of the most pervasive tools of engineering. Whenever the state of a system must be estimated from noisy sensor information, some kind of state estimator is employed to fuse the data from different sensors together to produce an accurate estimate of the true system state. When the system dynamics and observation models are linear, the minimum mean squared error (MMSE) estimate may be computed using the Kalman filter. However, in most applications of interest the system dynamics and observation equations are nonlinear and suitable extensions to the Kalman Filter have been sought like the extended kalman filters ,the unscented ones and the extended complex kalman filters

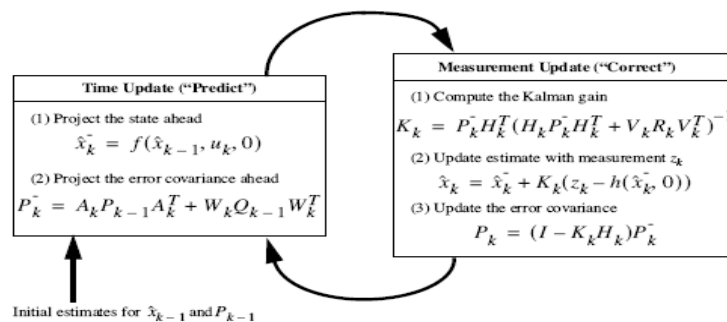
..

A BRIEF VIEW OF THE FILTERS

KALMAN FILTERS



EXTENDED KALMAN FILTERS



UNSCENTED KALMAN FILTERS

1. The set of sigma points are created to the augmented system

2. The transformed set is given by instantiating each point through the process model,

$$\mathcal{X}_i(k+1|k) = \mathbf{f}[\mathcal{X}_i^a(k|k), \mathbf{u}(k), k].$$

3. The predicted mean is computed as

$$\hat{\mathbf{x}}(k+1|k) = \sum_{i=0}^{2n^a} W_i \mathcal{X}_i^a(k+1|k).$$

4. And the predicted covariance is computed as

$$\mathbf{P}(k+1|k) = \sum_{i=0}^{2n^a} W_i \{\mathcal{X}_i(k+1|k) - \hat{\mathbf{x}}(k+1|k)\} \{\mathcal{X}_i(k+1|k) - \hat{\mathbf{x}}(k+1|k)\}^T$$

5. Instantiate each of the prediction points through the observation model,

$$\mathcal{Z}_i(k+1|k) = \mathbf{h}[\mathcal{X}_i(k+1|k), \mathbf{u}(k), k]$$

6. The predicted observation is calculated by

$$\hat{\mathbf{z}}(k+1|k) = \sum_{i=1}^{2n^a} W_i \mathcal{Z}_i(k+1|k).$$

7. Since the observation noise is additive and independent, the innovation covariance is

$$\mathbf{P}_{\nu\nu}(k+1|k) = \mathbf{R}(k+1) + \sum_{i=0}^{2n^a} W_i \{\mathcal{Z}_i(k|k-1) - \hat{\mathbf{z}}(k+1|k)\} \{\mathcal{Z}_i(k|k-1) - \hat{\mathbf{z}}(k+1|k)\}^T$$

8. Finally the cross correlation matrix is determined by

$$\mathbf{P}_{\mathbf{zx}}(k+1|k) = \sum_{i=0}^{2n^a} W_i \{\mathcal{X}_i(k|k-1) - \hat{\mathbf{x}}(k+1|k)\} \{\mathcal{Z}_i(k|k-1) - \hat{\mathbf{z}}(k+1|k)\}^T$$

EXTENDED COMPLEX

KALMAN FILTERS

A complex model has been developed to track a distorted signal that belongs to a power system. The model inherently takes care of the frequency measurement along with the amplitude and phase of the signals

CHAPTER 1

INTRODUCTION TO HARMONIC FREQUENCY ESTIMATION

INTRODUCTION TO HARMONIC FREQUENCY ESTIMATION

The frequency harmonic estimation in the power system has been a field of study lately due to an increasing numbers of high voltage transmission system having static var compensators placed at strategic location which continuously inject time varying harmonic in the system. This gives rise to non-stationary harmonic voltages and currents in the distribution system.

This paper presents an integrated approach to design an optimal power frequency estimator of harmonic components of a power network in the presence of frequency variations. This has led to the study of Kalman filter, Extended kalman filter and unscented kalman filter characteristic and a subsequent implementation of the study to design the optimal filter which was left to the forthcoming batches to implement. Due to the time factor we could simulate only the filter characteristic under various forms of signal and hence comprehend on the basis of the result which filters best suits to optimally design a filter.

The Kalman filter addresses the general problem of trying to estimate the state of a discrete-time controlled process that is governed by a linear stochastic difference equation. But what happens if the process to be estimated and the measurement relationship to the process is non-linear? Some of the most interesting and successful applications OF KALMAN filtering have been such situations. A Kalman filter that linearizes about the current mean and covariance is referred to as an extended Kalman filter or EKF.

In something a kin to a Taylor series, we can linearize the estimation around the current estimate using the partial derivatives of the process and measurement functions to compute estimates even in the face of non-linear relationships. Let us assume that our process again has a state vector, but that the process is now governed by the non-linear stochastic difference equations.

Commonly asked questions about the harmonics in the power system:

SALIENT FEATURES OF POWER SYSTEM HARMONICS

- Ideally, voltage and current waveforms are perfect sinusoids. However, because of the increased popularity of electronic and other non-linear loads, these waveforms quite often become distorted. This deviation from a perfect sine wave can be represented by harmonics—sinusoidal components having a frequency that is an integral multiple of the fundamental frequency. Thus, a pure voltage or current sine wave has no distortion and no harmonics, and a non-sinusoidal wave has distortion and harmonics. To quantify the distortion, the term total harmonic distortion (THD) is used. The term expresses the distortion as a percentage of the fundamental (pure sine) of voltage and current waveform.
- Current harmonics are a problem because they cause increased losses in the customer and utility power system components. Transformers are especially sensitive to this problem and may need to be derated to as much as 50% capacity when feeding loads with extremely distorted current waveforms (current total harmonic distortion above 100%). ANSI/IEEE C57.110-1986 (IEEE Recommended Practice for Establishing Transformer Capacity When Supplying Nonsinusoidal Load Currents) states that a transformer subject to nonsinusoidal load current having more than 5% total harmonics distortion needs to be derated. However, when current THD is below 15%, the derating of the transformer would be so small that it can be neglected. On the other hand, when current THD exceeds 15%, the transformer capability should be evaluated by professional using IEEE recommendations. It is important to clarify that these IEEE recommendations do not apply to transformers especially designed to feed nonsinusoidal loads. Underwriters Laboratories (UL) tests and rates this special type of transformer, also called a “K-factor” transformer. Loads with highly distorted current waveforms also have a very poor power factor; because of this, they use excessive power system capacity and could be a cause of overloading. Voltage source electronic adjustable speed drives (ASD) often have a total power factor of approximately 65% because of the highly distorted current. This total power factor could be corrected to approximately 85% using line-side chokes (reactors) on the drive. The chokes limit the rate of rise and the peak value of the line current, dramatically reducing the current THD. In addition, current harmonics can distort the voltage waveform and cause voltage harmonics. Voltage distortion affects not only sensitive electronic loads but also electric motors and capacitor banks. In electric motors, Negative sequence harmonics (i.e. 5th, 11th, 17th), so called because their sequence (ABC or ACB) is opposite of the fundamental sequence, produce rotating magnetic fields. These fields rotate in the opposite direction of the fundamental magnetic field and could cause not only overheating but also mechanical oscillations in the motor-load system.
- When non-linear loads are a considerable part of the total load in the facility (more than 20%), there is a chance of a harmonics problem. Another consideration is the amount of current distortion produced by the non-linear loads.

Electronic ballasts, for example, come with current THD ranging from 6% to 100%. It is important to avoid electronic ballasts with more than 20% current THD. PWM ASDs typically produce close to 100% current THD, which can be reduced to less than half by installing inexpensive 3% impedance line-side reactors (chokes). Another important way to check for harmonic currents is to measure the current in the neutral of a 3-phase 4-wire system. If the neutral current is considerably higher than the value predicted from the imbalance in the phase currents, there is a good possibility of heavy presence of triplen harmonics. Other signs of current harmonics include inexplicable higher-than-normal temperatures in the transformer, voltage distortion and high crest factor.

Harmonics have higher Frequency than the fundamental (60Hz); the 5th harmonic, or example, has a frequency of 300 HZ. 3^{Pulse} Width Modulation is the most common technology used in ASDs under 200HP.

- The crest factor of any waveform is the ratio of the peak value to the RMS value. In a perfect sine wave, the crest factor is 1.414. Crest factors different than 1.414 indicate distortion in the waveform. Typically distorted current waveforms have crest factors higher than 1.414, and distorted voltage waveforms have crest factors lower than 1.414. Distorted voltage waveforms with crest factors lower than 1.414 are called “flat top” voltage waveform voltage waveforms. The Computer and Business Equipment Manufacturers Association (CBEMA) recommend a method for derating transformers based on the current crest factor. CBEMA defines the transformer harmonic-derating factor (THDF) as the ratio of 1.414 to the transformer current crest factor. The derated KVA of the transformer would be the nominal KVA of the transformer multiplied by the THDF. This method, however, should only be applied when the distortion in the current is caused by single-phase, non-linear loads (like personal computers) and other office equipment. Also, when using this method, the crest factor measurements should always be made in the secondary of the transformer. A more complete method for transformer derating is the one described by ANSI/IEEE C57.110-1986.
- Harmonics only mean trouble if the power system is not designed to handle them. High harmonic neutral currents are a problem only if the neutral is not properly sized. Current harmonics are not a problem to a transformer if it is derated appropriately. Even some voltage distortion below 8% THD at the point of utilization) is acceptable as long as sensitive equipment is not affected. However, it is always important to be aware of the presence of harmonics and to try to minimize them by purchasing low distortion electronic ballasts and reactors for PWM ASDs. This will not only keep the harmonics in check and improve the power factor in the facility, but will also save energy by reducing losses on power system components. In addition, any time there is a considerable increase of non-linear loads, it is important to check power system components to prevent problems.

CHAPTER 2

COMPUTATIONAL ORIGIN OF THE FILTER

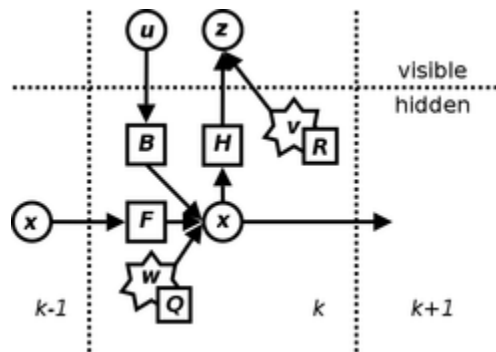
THE COMPUTATIONAL ORIGINS OF THE FILTER

To estimate a process with non-linear difference and measurement relationship, we begin by writing new governing equations that linearize an estimate, and then it involves the formulation of the measurement and update EQUATION. The subsequent steps involve the Measurement of Update equations which seem to have incremented the time variable t , so t in the Measurement Update equations is the same as $t+1$ in the Time Update equations. An important feature of the EKF is that the Jacobian in the equation for the Kalman gain serves to correctly propagate or "magnify" only the relevant component of the measurement information. For example, if there is not a one-to-one mapping between the measurement and the state then, the Jacobian affects the extended Kalman gain, so that it only magnifies the portion of the residual that does affect the state. Of over all measurements there is not a one-to-one mapping between the measurement and the state via $h(\cdot)$, then as you might expect the filter will quickly diverge. The control theory term to describe this situation is unobservable.

SO TO OVERCOME THE DISADVANTAGES OF EXTENDED KALMAN FILTER AS STATED ABOVE THE UNSCENTED KALMAN FILTER SEEMS THE MOST OPTIMAL FILTER TO DESIGN THE FREQUENCY ESTIMATOR. HOWEVER IF THE SYSTEM NOISE IS LESS EXTENDED KALMAN FILTER CAN ALSO BE USED TO A GREATER DEGREE OF ACCURACY.

DYNAMIC SYSTEM MODEL

Kalman filters are based on linear dynamical systems discretised in the time domain. They are modelled on a Markov chain built on linear operators perturbed by Gaussian noise. The state of the system is represented as a vector of real numbers. At each discrete time increment, a linear operator is applied to the state to generate the new state, with some noise mixed in, and optionally some information from the controls on the system if they are known. Then, another linear operator mixed with more noise generates the visible outputs from the hidden state. The Kalman filter may be regarded as analogous to the hidden Markov model, with the key difference that the hidden state variables are continuous (as opposed to being discrete in the hidden Markov model). Additionally, the hidden Markov model can represent an arbitrary distribution for the next value of the state variables, in contrast to the Gaussian noise model that is used for the Kalman filter. There is a strong duality between the equations of the Kalman Filter and those of the hidden Markov model. In order to use the Kalman filter to estimate the internal state of a process given only a sequence of noisy observations, one must model the process in accordance with the framework of the Kalman filter. This means specifying the matrices F_k , H_k , Q_k , R_k , and sometimes B_k for each time-step k as described below.



. Model underlying the kalman filter

(Circles are vectors, squares are matrices, and stars represent Gaussian noise with the associated covariance matrix at the lower right.)

The Kalman filter model assumes the true state at time k is evolved from the state at $(k - 1)$ according to

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k$$

where

- \mathbf{F}_k is the state transition model which is applied to the previous state \mathbf{x}_{k-1} ;
- \mathbf{B}_k is the control-input model which is applied to the control vector \mathbf{u}_k ;
- \mathbf{w}_k is the process noise which is assumed to be drawn from a zero mean multivariate normal distribution with covariance \mathbf{Q}_k .

$$\mathbf{w}_k \sim N(0, \mathbf{Q}_k)$$

At time k an observation (or measurement) \mathbf{z}_k of the true state \mathbf{x}_k is made according to

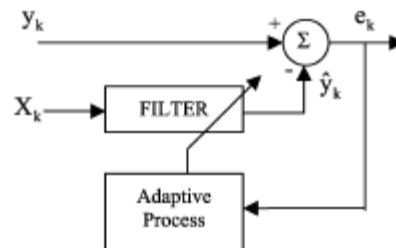
$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k$$

where \mathbf{H}_k is the observation model which maps the true state space into the observed space and \mathbf{v}_k is the observation noise which is assumed to be zero mean Gaussian white noise with covariance \mathbf{R}_k .

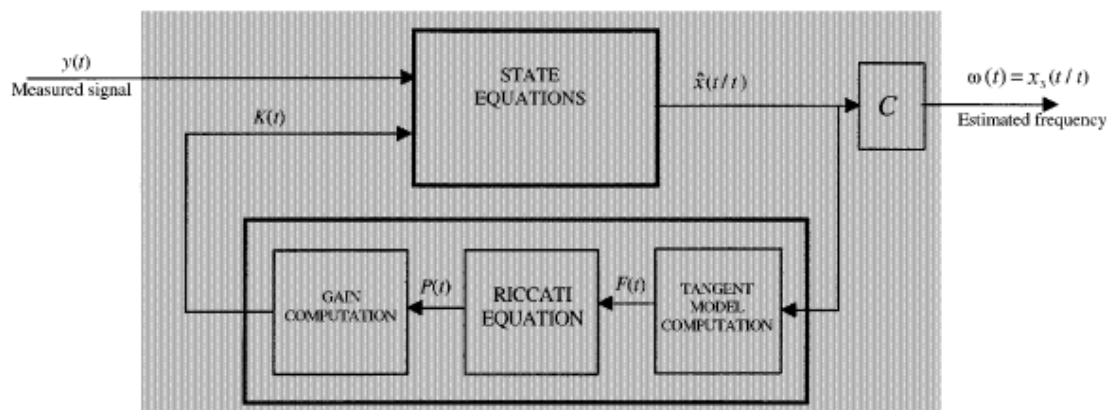
$$\mathbf{v}_k \sim N(0, \mathbf{R}_k)$$

The initial state, and the noise vectors at each step $\{\mathbf{x}_0, \mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_1, \dots, \mathbf{v}_k\}$ are all assumed to be mutually independent.

Many real dynamical systems do not exactly fit this model; however, because the Kalman filter is designed to operate in the presence of noise, an approximate fit is often good enough for the filter to be very useful. Variations on the Kalman filter described below allow richer and more sophisticated models.



BASIC ADAPTIVE FILTER



EXTENDED KALMAN FILTER FREQUENCY TRACKER

CHAPTER 3

KALMAN FILTER

The kalman filter is a recursive estimator. This means that only the estimated state from the previous time step and the current measurement are needed to compute the estimate for the current state. The state of the filter is represented by two variable matrices

$\hat{\mathbf{x}}_{k|k}$:- The estimate of the state at time K and

$\mathbf{P}_{k|k}$:- The error covariance matrix (A measure of the measured accuracy of the state).

The Process to be Estimated

The Kalman filter addresses the general problem of trying to estimate the state of a discrete-time controlled process that is governed by the linear stochastic difference equation:-

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}u_k + \mathbf{w}_{k-1},$$

With a measurement: -

$$z_k = \mathbf{H}\mathbf{x}_k + v_k.$$

The random variables represent the process and measurement noise (respectively). They are assumed to be independent (of each other), white, and with normal probability distributions. The process noise covariance Q and measurement noise covariance R matrices might change with each time step or measurement, however here we assume they are constant matrix in the difference equation which relates the state at the previous time step to the state at the current step.

The Computational Origins of the Filter

We define $\hat{\mathbf{x}}_k^- \in \mathbb{R}^n$ (“super minus”) to be our **a priori state estimate at step k** given knowledge of the process prior to step k, and to be our **a posteriori state estimate at step k** given measurement. We can then define a priori and a posteriori estimate errors as

$$\begin{aligned} \mathbf{e}_k^- &= \mathbf{x}_k - \hat{\mathbf{x}}_k^-, \text{ and} \\ \mathbf{e}_k &= \mathbf{x}_k - \hat{\mathbf{x}}_k. \end{aligned}$$

The a priori estimate error covariance is then

$$\mathbf{P}_k^- = E[\mathbf{e}_k^- \mathbf{e}_k^{-T}],$$

And the a posteriori estimate error covariance is

$$\mathbf{P}_k = E[\mathbf{e}_k \mathbf{e}_k^T].$$

”

The next step involves finding an equation that computes an a posteriori state estimate as a linear combination of an a priori estimate and a weighted difference between an actual measurement and a measurement prediction

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-)$$

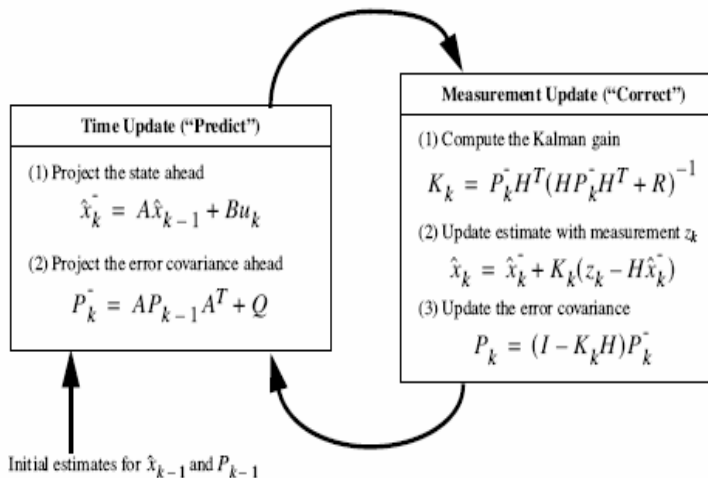
The kalman gain can be calculate from the equation

$$\begin{aligned} K_k &= P_k^- H^T (H P_k^- H^T + R)^{-1} \\ &= \frac{P_k^- H^T}{H P_k^- H^T + R} \end{aligned}$$

We see that as the R , measurement error covariance approaches zero, the gain K weights the residual more heavily. It is similar to that as the measurement error covariance approaches zero, the actual measurement is “trusted” more and more, while the predicted measurement is trusted less and less. On the other hand, as the a priori estimate error covariance approaches zero the actual measurement is trusted less and less, while the predicted measurement is trusted more and more

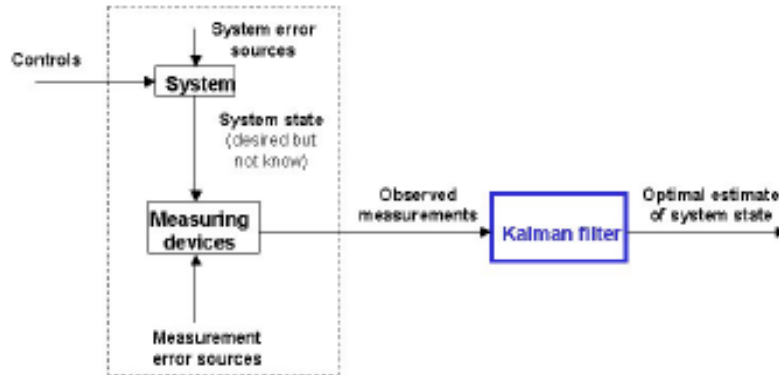
KALMAN FILTERING ALGORITHM

The Kalman filter estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements. As such, the equations for the Kalman filter fall into two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimates for the next time step. The measurement update equations are responsible for the feedback—i.e. for incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate



KALMAN FILTER PREDICTION ESTIMATION CYCLE

KALMAN FILTER



KALMAN FILTERING MODEL

The Kalman filter is a recursive estimator. This means that only the estimated state from the previous time step and the current measurement are needed to compute the estimate for the current state. In contrast to batch estimation techniques, no history of observations and/or estimates is required. It is unusual in being purely a time domain filter; most filters (for example, a low-pass filter) are formulated in the frequency domain and then transformed back to the time domain for implementation.

The state of the filter is represented by two variables:

- $\hat{\mathbf{x}}_{k|k}$, the estimate of the state at time k ;
- $\mathbf{P}_{k|k}$, the error covariance matrix (a measure of the estimated accuracy of the state estimate).

The Kalman filter has two distinct phases: Predict and Update. The predict phase uses the estimate from the previous timestep to produce an estimate of the current state. In the update phase, measurement information from the current timestep is used to refine this prediction to arrive at a new, (hopefully) more accurate estimate.

PREDICT

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_k \mathbf{u}_k \text{ (PREDICTED STATE)}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_k \text{ (PREDICTED ESTIMATE COVARIANCE)}$$

UPDATE

$$\tilde{\mathbf{y}}_k = \mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \text{ (INNOVATION OR MEASUREMENT RESIDUAL)}$$

$$\begin{aligned}
\mathbf{S}_k &= \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \text{ (INNOVATION (OR RESIDUAL) COVARIANCE)} \\
\mathbf{K}_k &= \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1} \text{ (OPTIMAL KALMAN GAIN)} \\
\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k \text{ (UPDATED STATE ESTIMATE)} \\
\mathbf{P}_{k|k} &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \text{ (UPDATED ESTIMATE COVARIANCE)}
\end{aligned}$$

The formula for the updated estimate covariance above is only valid for the optimal Kalman gain. Usage of other gain values require a more complex formula found in the *derivations* section.

INVARIANTS

If the model is accurate, and the values for $\hat{\mathbf{x}}_{0|0}$ and $\mathbf{P}_{0|0}$ accurately reflect the distribution of the initial state values, then the following invariants are preserved: all estimates have mean error zero

- $\mathbf{E}[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}] = \mathbf{E}[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}] = \mathbf{0}$
- $\mathbf{E}[\tilde{\mathbf{y}}_k] = \mathbf{0}$

and covariance matrices accurately reflect the covariance of estimates

- $\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})$
- $\mathbf{P}_{k|k-1} = \text{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})$
- $\mathbf{S}_k = \text{cov}(\tilde{\mathbf{y}}_k)$

DERIVATIONS

Deriving the posterior estimate covariance matrix

Starting with our invariant on the error covariance $\mathbf{P}_{k|k}$ as above

$$\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})$$

substitute in the definition of $\hat{\mathbf{x}}_{k|k}$

$$\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - (\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k))$$

and substitute $\tilde{\mathbf{y}}_k$

$$\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - (\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1})))$$

and \mathbf{z}_k

$$\mathbf{P}_{k|k} = \text{cov}(\mathbf{x}_k - (\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1})))$$

and by collecting the error vectors we get

$$\mathbf{P}_{k|k} = \text{cov}((\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) - \mathbf{K}_k \mathbf{v}_k)$$

Since the measurement error \mathbf{v}_k is uncorrelated with the other terms, this becomes

$$\mathbf{P}_{k|k} = \text{cov}((\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})) + \text{cov}(\mathbf{K}_k \mathbf{v}_k)$$

by the properties of vector covariance this becomes

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \text{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \text{cov}(\mathbf{v}_k) \mathbf{K}_k^T$$

which, using our invariant on $\mathbf{P}_{k|k-1}$ and the definition of \mathbf{R}_k becomes

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T$$

This formula is valid no matter what the value of \mathbf{K}_k . It turns out that if \mathbf{K}_k is the optimal Kalman gain, this can be simplified further as shown below.

KALMAN GAIN DERIVATION

The Kalman filter is a minimum mean-square error estimator. The error in the posterior state estimation is

$$\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}$$

We seek to minimize the expected value of the square of the magnitude of this vector, $\mathbb{E}[|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}|^2]$. This is equivalent to minimizing the trace of the posterior estimate covariance matrix $\mathbf{P}_{k|k}$. By expanding out the terms in the equation above and collecting, we get:

$$\begin{aligned} \mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k) \mathbf{K}_k^T \\ &= \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T \end{aligned}$$

The trace is minimized when the matrix derivative is zero:

$$\frac{\partial \text{tr}(\mathbf{P}_{k|k})}{\partial \mathbf{K}_k} = -2(\mathbf{H}_k \mathbf{P}_{k|k-1})^T + 2\mathbf{K}_k \mathbf{S}_k = 0$$

Solving this for \mathbf{K}_k yields the Kalman gain:

$$\begin{aligned}\mathbf{K}_k \mathbf{S}_k &= (\mathbf{H}_k \mathbf{P}_{k|k-1})^T = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \\ \mathbf{K}_k &= \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1}\end{aligned}$$

This gain, which is known as the *optimal Kalman gain*, is the one that yields MMSE estimates when used.

Simplification of the posterior error covariance formula

The formula used to calculate the posterior error covariance can be simplified when the Kalman gain equals the optimal value derived above. Multiplying both sides of our Kalman gain formula on the right by $\mathbf{S}_k \mathbf{K}_k^T$, it follows that

$$\mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{K}_k^T$$

Referring back to our expanded formula for the posterior error covariance,

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T$$

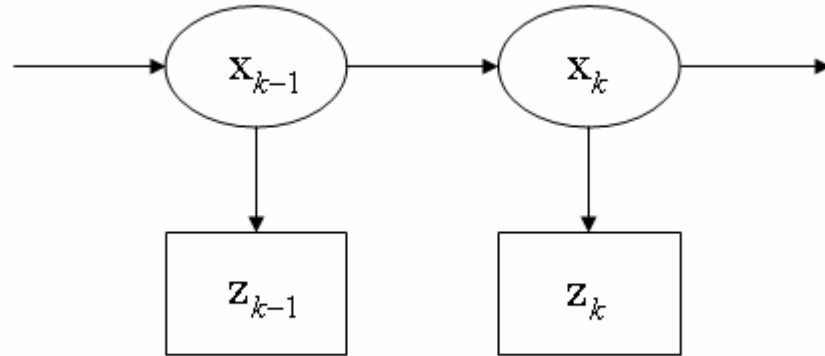
we find the last two terms cancel out, giving

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}.$$

This formula is computationally cheaper and thus nearly always used in practice, but is only correct for the optimal gain. If arithmetic precision is unusually low causing problems with numerical stability, or if a non-optimal Kalman gain is deliberately used, this simplification cannot be applied; the posterior error covariance formula as derived above must be used.

RELATIONSHIP TO RECURSIVE BAYESIAN ESTIMATION

The true state is assumed to be an unobserved Markov process, and the



measurements are the

Because of the Markov assumption, the true state is conditionally independent of all earlier states given the immediately previous state.

$$p(\mathbf{x}_k | \mathbf{x}_0, \dots, \mathbf{x}_{k-1}) = p(\mathbf{x}_k | \mathbf{x}_{k-1})$$

Similarly the measurement at the k -th timestep is dependent only upon the current state and is conditionally independent of all other states given the current state.

$$p(\mathbf{z}_k | \mathbf{x}_0, \dots, \mathbf{x}_k) = p(\mathbf{z}_k | \mathbf{x}_k)$$

Using these assumptions the probability distribution over all states of the HMM can be written simply as:

$$p(\mathbf{x}_0, \dots, \mathbf{x}_k, \mathbf{z}_1, \dots, \mathbf{z}_k) = p(\mathbf{x}_0) \prod_{i=1}^k p(\mathbf{z}_i | \mathbf{x}_i) p(\mathbf{x}_i | \mathbf{x}_{i-1})$$

However, when the Kalman filter is used to estimate the state \mathbf{x} , the probability distribution of interest is that associated with the current states conditioned on the measurements up to the current timestep. (This is achieved by marginalizing out the previous states and dividing by the probability of the measurement set.)

This leads to the *predict* and *update* steps of the Kalman filter written probabilistically. The probability distribution associated with the predicted state is product of the probability distribution associated with the transition from the $(k-1)$ th timestep to the k th and the probability distribution associated with the previous state, with the true state at $(k-1)$ integrated out.

$$p(\mathbf{x}_k | \mathbf{Z}_{k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{Z}_{k-1}) d\mathbf{x}_{k-1}$$

The measurement set up to time t is

$$\mathbf{Z}_t = \{\mathbf{z}_1, \dots, \mathbf{z}_t\}$$

The probability distribution of updated is proportional to the product of the measurement likelihood and the predicted state.

$$p(\mathbf{x}_k | \mathbf{Z}_k) = \frac{p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{Z}_{k-1})}{p(\mathbf{z}_k | \mathbf{Z}_{k-1})}$$

The denominator

$$p(\mathbf{z}_k | \mathbf{Z}_{k-1}) = \int p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{Z}_{k-1}) d\mathbf{x}_k$$

is an unimportant normalization term.

The remaining probability density functions are

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{x}_{k-1}) &= N(\mathbf{F}_k \mathbf{x}_{k-1}, \mathbf{Q}_k) \\ p(\mathbf{z}_k | \mathbf{x}_k) &= N(\mathbf{H}_k \mathbf{x}_k, \mathbf{R}_k) \\ p(\mathbf{x}_{k-1} | \mathbf{Z}_{k-1}) &= N(\hat{\mathbf{x}}_{k-1}, \mathbf{P}_{k-1}) \end{aligned}$$

Note that the PDF at the previous timestep is inductively assumed to be the estimated state and covariance. This is justified because, as an optimal estimator, the Kalman filter makes best use of the measurements, therefore the PDF for \mathbf{x}_k given the measurements \mathbf{Z}_k is the Kalman filter estimate.

INFORMATION FILTER

In the information filter, or inverse covariance filter, the estimated covariance and estimated state are replaced by the information matrix and information vector respectively.

$$\begin{aligned} \mathbf{Y}_{k|k} &\equiv \mathbf{P}_{k|k}^{-1} \\ \hat{\mathbf{y}}_{k|k} &\equiv \mathbf{P}_{k|k}^{-1} \hat{\mathbf{x}}_{k|k} \end{aligned}$$

Similarly the predicted covariance and state have equivalent information forms,

$$\mathbf{Y}_{k|k-1} \equiv \mathbf{P}_{k|k-1}^{-1}$$

$$\hat{\mathbf{y}}_{k|k-1} \equiv \mathbf{P}_{k|k-1}^{-1} \hat{\mathbf{x}}_{k|k-1}$$

as have the measurement covariance and measurement vector.

$$\mathbf{I}_k \equiv \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k$$

$$\mathbf{i}_k \equiv \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{z}_k$$

The information update now becomes a trivial sum.

$$\mathbf{Y}_{k|k} = \mathbf{Y}_{k|k-1} + \mathbf{I}_k$$

$$\hat{\mathbf{y}}_{k|k} = \hat{\mathbf{y}}_{k|k-1} + \mathbf{i}_k$$

The main advantage of the information filter is that N measurements can be filtered at each timestep simply by summing their information matrices and vectors.

$$\mathbf{Y}_{k|k} = \mathbf{Y}_{k|k-1} + \sum_{j=1}^N \mathbf{I}_{k,j}$$

$$\hat{\mathbf{y}}_{k|k} = \hat{\mathbf{y}}_{k|k-1} + \sum_{j=1}^N \mathbf{i}_{k,j}$$

To predict the information filter the information matrix and vector can be converted back to their state space equivalents, or alternatively the information space prediction can be used.

$$\mathbf{M}_k = [\mathbf{F}_k^{-1}]^T \mathbf{Y}_{k|k} \mathbf{F}_k^{-1}$$

$$\mathbf{C}_k = \mathbf{M}_k [\mathbf{M}_k + \mathbf{Q}_k^{-1}]^{-1}$$

$$\mathbf{L}_k = \mathbf{I} - \mathbf{C}_k$$

$$\mathbf{Y}_{k|k-1} = \mathbf{L}_k \mathbf{M}_k \mathbf{L}_k^T + \mathbf{C}_k \mathbf{Q}_k^{-1} \mathbf{C}_k^T$$

$$\hat{\mathbf{y}}_{k|k-1} = \mathbf{L}_k [\mathbf{F}_k^{-1}]^T \hat{\mathbf{y}}_{k|k}$$

Note that if F and Q are time invariant these values can be cached. Note also that F and Q need to be invertible.

KALMAN FILTERS

To use a standard Kalman filter to estimate a signal, the process we're measuring has to be describable by linear system equations. A *linear system* is a process that can be described by the following two equations:

$$\text{State equation: } x_{k+1} = Ax_k + Bu_k + w_k$$

(1)

$$\text{Output Equation: } y_k = Cx_k + v_k$$

(2)

These equations define a linear system because they don't contain any exponential functions, trigonometric functions, or any other functions that wouldn't make a straight line when plotted on a graph. Equations 1 and 2 have several variables:

- A , B , and C are matrices
- k is the time index
- x is called the state of the system
- u is a known input to the system (called the control signal)
- y is the measured output
- w and v are the noise-- w is called the process noise, and v is called the measurement noise

Other than the time index, each of these variables is (in general) vectors and therefore contains more than one element.

In state-estimation problems, we want to estimate x because it contains all of the information about the system. The problem is, we can't measure x directly. Instead we measure y , which is a function of x that's corrupted by the noise of v . We can use y to help us obtain an estimate of x , but we can't necessarily take the information from y at its face value because it's corrupted by noise.

As an example

Suppose that our system is a tank, a mobile robot, a pizza-delivery car, or some other vehicle moving in a straight line. We can say that the state consists of the vehicle position and velocity. The input u is the acceleration and the output y is the measured position. Let's further suppose that we're able to measure the position every T seconds. This system can be described thus:

$$\begin{aligned}x_{k+1} &= \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} T^2/2 \\ T \end{bmatrix} u_k + w_k \\ y_k &= \begin{bmatrix} 1 & 0 \end{bmatrix} x_k + v_k\end{aligned}$$

(3)

x_k is a vector that contains vehicle position and velocity at time k , u_k is a scalar that is equal to the acceleration, and y_k is a scalar that is equal to the measured position. w_k is a vector that has process noise due to potholes, uncertainties in our knowledge of u_k , and other unmodeled effects. Finally, v_k is a scalar that's equal to the measurement noise (that is, instrumentation error).

Now suppose that we want to control the vehicle to follow a specific path, or we want to estimate the vehicle position for some other reason. We could just use y_k as our position estimate, but y_k is noisy. We could do better by using a Kalman filter. This is because a Kalman filter uses not only the position measurement y_k , but also the information contained in the state equation. The Kalman filter equations can be written like this:

$$\begin{aligned}K_k &= P_k C^T (C P_k C^T + R)^{-1} \\ \hat{x}_{k+1} &= (A \hat{x}_k + B u_k) + K_k (y_k - C \hat{x}_k) \\ P_{k+1} &= A(I - K_k C) P_k A^T + Q\end{aligned}$$

(4)

where the time step $k = 0, 1, 2, \dots$. Once again, the Kalman filter is called a *linear* filter because the \hat{x} equation doesn't contain any exponential functions, trigonometric functions, or any other functions that would not appear as a straight line on a graph. \hat{x}_k is the estimate of x_k

- K_k is called the Kalman gain (it is a matrix)
- P_k is called the estimation-error covariance (also a matrix)
- Q is the covariance of the process noise w_k , and R is the covariance of the measurement noise v_k (two more matrices)
- The -1 superscript indicates matrix inversion
- The T superscript indicates matrix transposition
- I is the identity matrix

In order to initialize the Kalman filter, we need to start with an estimate \hat{x}_0 of the state at the initial time. We also need to start with an initial estimation error covariance P_0 , which represents our uncertainty in our initial state estimate. If we're very confident in our initial estimate \hat{x}_0 , then P_0 should be small. If we're very uncertain of our initial estimate \hat{x}_0 , then P_0 should be large. In the long run, these initialization values won't make much difference to the filter's performance.

Linearity limitations

The Kalman filter is a linear filter that can be applied to a linear system. Unfortunately, linear systems don't really exist--all systems are ultimately nonlinear. Even the simple $I = V/R$ relationship of Ohm's law is only an approximation over a limited range. If the voltage across a resistor exceeds a certain value, Ohm's law breaks down. Figure 1 shows a typical relationship between the current through a resistor and the voltage across the resistor. At small input voltages the relationship is a straight line, but if the power dissipated by the resistor exceeds some value, the relationship becomes very nonlinear. Even a device as simple as a resistor is only approximately linear, and even then only in a limited range of operation. This illustrates the fact that linear systems do not exist in the real world.

However, many systems are close enough that linear estimations (for example, standard Kalman filters) give good results. But "close enough" can only be carried so far. Eventually we'll run across a system that doesn't behave linearly even over a small range of operation, and the standard Kalman filter no longer gives good results. In these cases, we'll need to explore nonlinear filters.

Nonlinear filtering can be difficult and complex; it's certainly not as well understood as linear filtering. However, some nonlinear estimation methods have become (or are becoming) widespread. These include nonlinear extensions of the Kalman filter, **"unscented" filtering, and particle filtering**.

The standard Kalman filter doesn't directly apply to nonlinear systems. However, if we *linearize* a nonlinear system we can use linear-estimation methods to estimate the states. In order to linearize a nonlinear system various mathematical tool such as Taylor series expansion etc are used.

TAYLOR SERIES EXPANSION:-

The key to nonlinear Kalman filtering is to expand the nonlinear terms of the system equation in a Taylor series expansion around a nominal point. A Taylor series expansion of a nonlinear function can be written as:

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(\bar{x}) \Delta x^n}{n!} \\ &= f(\bar{x}) + f'(\bar{x}) \Delta x + \\ &\quad f''(\bar{x}) \Delta x^2 / 2 + \dots \end{aligned}$$

(5)

In Equation 5:

- $\Delta x = x - \bar{x}$
- $f^{(n)}(\bar{x})$ is the n th derivative of $f(x)$, evaluated at $x = \bar{x}$

Taking an example. Suppose we want to expand $f(x) = \cos(x)$ in a Taylor series around the point $\bar{x} = 0$. Remember that the derivative of $\cos(x)$ is $-\sin(x)$, and the derivative of $\sin(x)$ is $\cos(x)$. That means we can write the Taylor series expansion of $\cos(x)$ as:

$$\cos(x) = \cos(\bar{x}) - \sin(\bar{x})\Delta x - \cos(\bar{x})\Delta x^2 / 2 + \dots$$

(6)

Since we're expanding $\cos(x)$ around the nominal point $x = 0$, we see that $\bar{x} = 0$, and $\Delta x = x - \bar{x} = x$. The Taylor series expansion of $\cos(x)$ becomes equal to:

$$\begin{aligned} \cos(x) &= \cos(0) - \sin(0)x - \cos(0)x^2 / 2 + \dots \\ &= 1 - x^2 / 2 + \dots \end{aligned}$$

(7)

If we use a "second-order" Taylor series expansion of $\cos(x)$, we can say that $\cos(x)$ is approximately equal to $1 - x^2 / 2$. (It's called "second-order" because it includes terms up to and including the second power of x .) In other words, we can ignore the rest of the terms in the Taylor series. This is because additional terms in the Taylor series involve higher powers of x that are divided by ever-increasing factorials. If x is small, as we raise x to higher and higher powers and divide by larger and larger factorials, the additional terms in the Taylor series become insignificant compared with the lower order terms. Table 1 shows $\cos(x)$ and its second-order Taylor series expansion for various values of x . We see that as x gets smaller (that is, as it gets closer to the nominal point $\bar{x} = 0$), the Taylor series expansion gives a better approximation to the true value of $\cos(x)$.

The second-order Taylor series expansion of $\cos(x)$ is accurate for small values of x , but it quickly becomes inaccurate as x gets large (in other words, as x gets farther away from the Taylor series expansion point)

x	$\cos(x)$	$1 - x^2 / 2$
0	1	1
0.25	0.969	0.969
0.5	0.878	0.875
0.75	0.732	0.719
1.00	0.540	0.500
1.25	0.315	0.219
1.50	0.071	-0.125

TABLE 6.1 Taylor series expansion of $\cos x$

Linearizing a function means expanding it in a "first-order" Taylor series around some expansion point. In other words, the first-order Taylor series expansion of a function $f(x)$ is equal to:

$$f(x) = f(\bar{x}) + f'(\bar{x})\Delta x$$

(8)

Figure 2 shows the function $\sin(x)$ along with its first-order Taylor series expansion around the nominal point $\bar{x} = 0$. Note that for small values of x , the two lines in the figure are quite close, which shows that the Taylor series expansion is a good approximation to $\sin(x)$. But as x gets larger the two lines diverge, so for large values of x , the Taylor series expansion is a poor approximation.

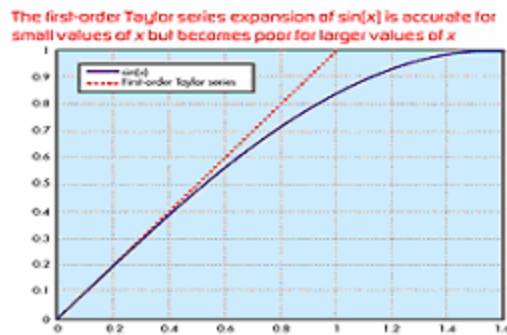


FIG 6.2 Taylor series expansion of $\sin x$

USING NONLINEAR KALMAN FILTERING TO ESTIMATE SIGNALS

The *Kalman filter* is a tool that estimates the variables of a wide range of processes. In mathematical terms we'd say that a Kalman filter estimates the states of a linear system. There are two reasons one might want to know the states of a system, whether linear or nonlinear:

- First, one might need to estimate states in order to control the system. For example, the electrical engineer needs to estimate the winding currents of a motor in order to control its position. The aerospace engineer needs to estimate the velocity of a satellite in order to control its orbit. The biomedical engineer needs to estimate blood-sugar levels in order to regulate insulin injection rates.
- Second, you might need to estimate system states because they're interesting in their own right. For example, the electrical engineer needs to estimate power-system parameters in order to predict failure probabilities. The aerospace engineer needs to estimate satellite position in order to intelligently schedule future satellite activities. The biomedical engineer needs to estimate blood-protein levels to evaluate the health of a patient.

The standard Kalman filter is an effective tool for estimation, but it's limited to linear systems. Most real-world systems are nonlinear, in which case Kalman filters don't directly apply. In the real world, nonlinear filters are used more often than linear filters because real systems are nonlinear. In fact, the very first use of Kalman filters involved nonlinear Kalman filters in NASA's space program in the 1960s.

SO LET US BEGIN WITH THE DISCUSSION OF KALMAN FILTERING PROCESS:

CHAPTER 4

EXTENDED KALMAN FILTER

EXTENDED KALMAN FILTER

When the nonlinearity of functions increases, the difference between the estimate and minimum position of the cost function increases. In the case of the cost function of EKF, past measurements are not ACCUMULATED; instead information on the cost function is stored as the estimate. Therefore, as the error of the estimate increases, the cost function itself contains an error. If estimation is repeated in this status, the error gradually accumulates. Even if new measurements are acquired, information on the cost function loses accuracy, influenced by the estimate that has accumulated the error thus far.

STATE-ESTIMATION

The basic framework for the EKF involves estimation of the state of a discrete-time nonlinear dynamic system,

$$\mathbf{x}_{k+1} = F(\mathbf{x}_k, \mathbf{v}_k) \quad (1)$$

$$\mathbf{y}_k = H(\mathbf{x}_k, \mathbf{n}_k), \quad (2)$$

Where \mathbf{x}_k represent the unobserved state of the system and \mathbf{y}_k is the only Observed signal. The *process* noise \mathbf{v}_k drives the dynamic system, and the *observation* noise is given by \mathbf{n}_k . Note that we are not assuming additively of the noise sources. The system dynamic model F and H are assumed known.

In state-estimation, the EKF is the standard method of choice to achieve a recursive (approximate) maximum-likelihood estimation of the state \mathbf{x}_k .

Parameter Estimation

The classic problem involves determining a nonlinear mapping

$$\mathbf{y}_k = G(\mathbf{x}_k, \mathbf{w}) \quad (3)$$

where \mathbf{x}_k is the input, \mathbf{y}_k is the output, and the nonlinear map G is parameterized by the vector \mathbf{w} . The nonlinear map, for example, may be a feed forward or recurrent neural network (\mathbf{w} are the weights), with numerous applications in regression, classification, and dynamic modeling. Learning corresponds to estimating the parameters \mathbf{w} . Typically, a training set is provided with

sample pairs consisting of known input and desired outputs, $\{\mathbf{x}_k, \mathbf{d}_k\}$. The error of the machine is defined as $\mathbf{e}_k = \mathbf{d}_k - G(\mathbf{x}_k, \mathbf{w})$, and the goal of learning involves solving for the parameters \mathbf{w} in order to minimize the expected squared error.

While a number of optimization approaches exist the EKF may be used to estimate the parameters by writing a new state-space representation

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \mathbf{u}_k \quad (4)$$

$$\mathbf{y}_k = G(\mathbf{x}_k, \mathbf{w}_k) + e_k. \quad (5)$$

Where the parameters \mathbf{w}_k correspond to a stationary process with identity state transition matrix, driven by process noise \mathbf{u}_k (the choice of variance determines tracking performance). The output \mathbf{y}_k corresponds to a nonlinear observation on \mathbf{w}_k . The EKF can then be applied directly as an efficient "second-order" technique for learning the parameters. In the linear case, the relationship between the Kalman Filter (KF) and Recursive Least Squares (RLS) in equation no : 3.

Dual Estimation

A special case arises when the input \mathbf{x}_k is unobserved, and requires coupling both state-estimation and parameter estimation. For these *dual estimation* problems, we again consider a discrete-time nonlinear dynamic system,

$$\mathbf{x}_{k+1} = F(\mathbf{x}_k, \mathbf{v}_k, \mathbf{w}) \quad (6)$$

$$\mathbf{y}_k = H(\mathbf{x}_k, \mathbf{n}_k, \mathbf{w}). \quad (7)$$

where both the system states \mathbf{x}_k and the set of model parameters \mathbf{w} for the dynamic system must be simultaneously estimated from only the observed noisy signal \mathbf{y}_k .

EXTENDED KALMAN FILTER

In the extended Kalman filter (EKF) the state transition and observation models need not be linear functions of the state but may instead be (differentiable) functions.

$$\begin{aligned}\mathbf{x}_k &= f(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{w}_k) \\ \mathbf{z}_k &= h(\mathbf{x}_k, \mathbf{v}_k)\end{aligned}$$

The function f can be used to compute the predicted state from the previous estimate and similarly the function h can be used to compute the predicted measurement from the predicted state. However, f and h cannot be applied to the covariance directly. Instead a matrix of partial derivatives (the Jacobian) is computed.

At each timestep the Jacobian is evaluated with current predicted states. These matrices can be used in the Kalman filter equations. This process essentially linearizes the non-linear function around the current estimate. This results in the following extended Kalman filter equations:

PREDICT

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= f(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_k, 0) \\ \mathbf{P}_{k|k-1} &= \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_k\end{aligned}$$

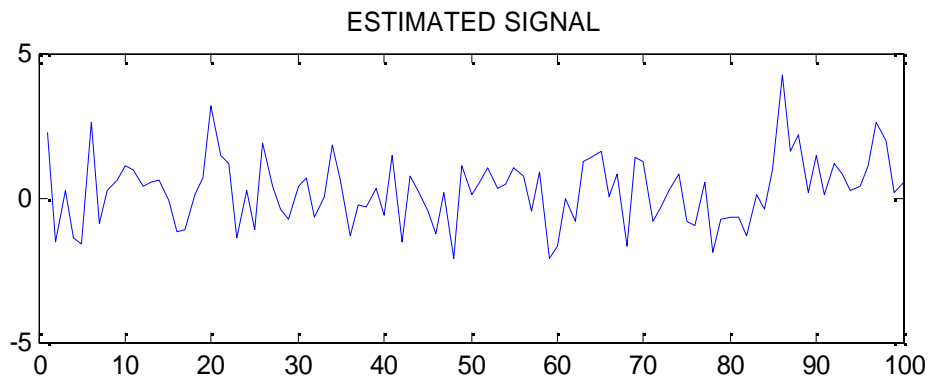
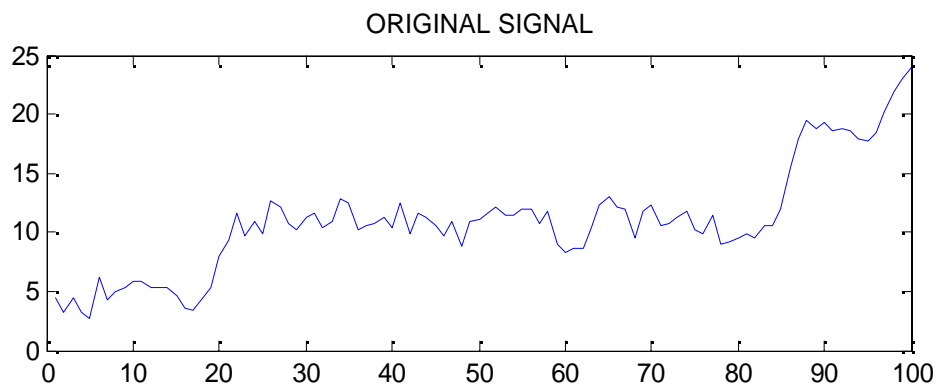
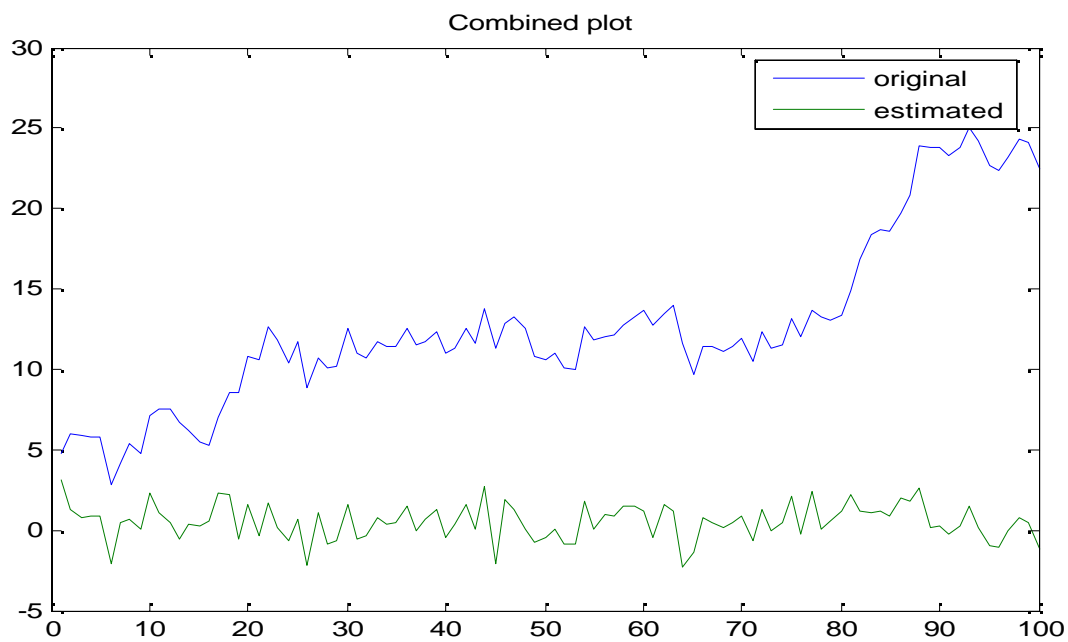
UPDATE

$$\begin{aligned}\tilde{\mathbf{y}}_k &= \mathbf{z}_k - h(\hat{\mathbf{x}}_{k|k-1}, 0) \\ \mathbf{S}_k &= \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \\ \mathbf{K}_k &= \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k \\ \mathbf{P}_{k|k} &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}\end{aligned}$$

Where the state transition and observation matrices are defined to be the following Jacobians

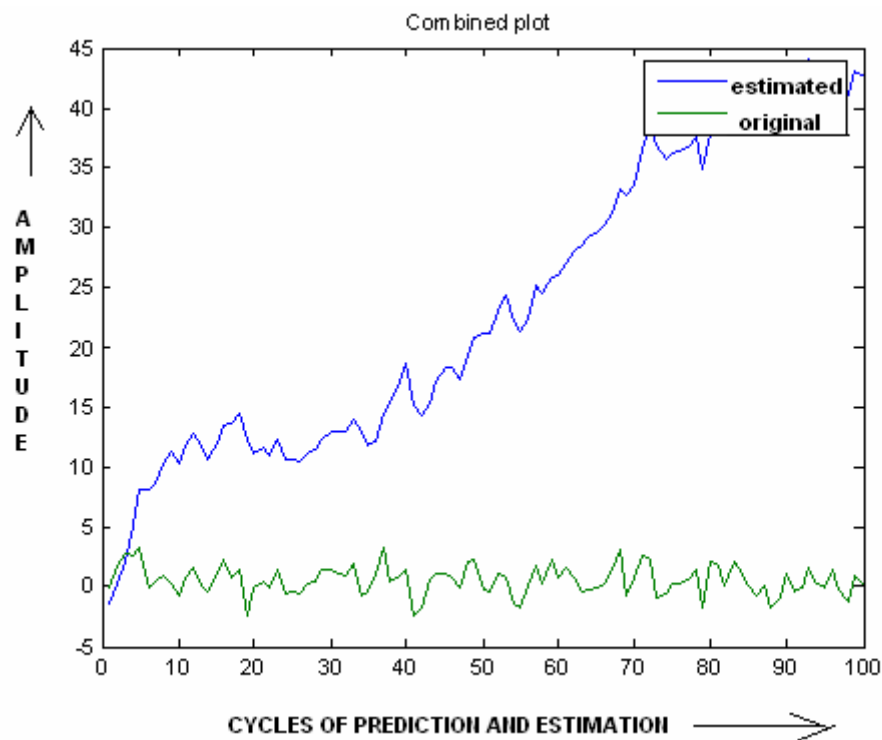
$$\mathbf{F}_k = \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_k} \quad \mathbf{H}_k = \left. \frac{\partial h}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k|k-1}}$$

EXTENDED KALMAN FILTER SIMULATION OUTPUT

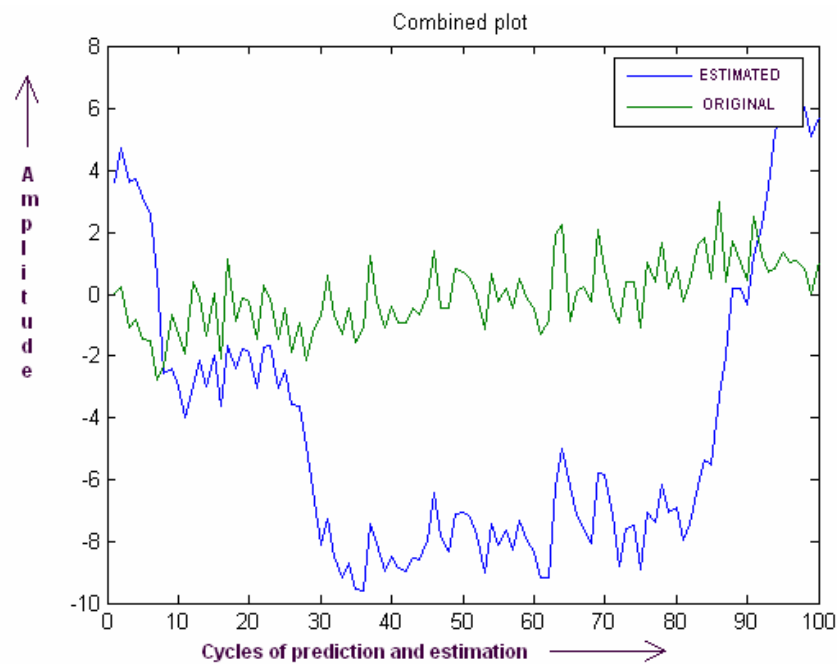


EXTENDED KALMAN FILTER ESTIMATE SIMULATION

LIMITATION OF EXTENDED KALMAN FILTER TO HIGH NONLINEARITIES

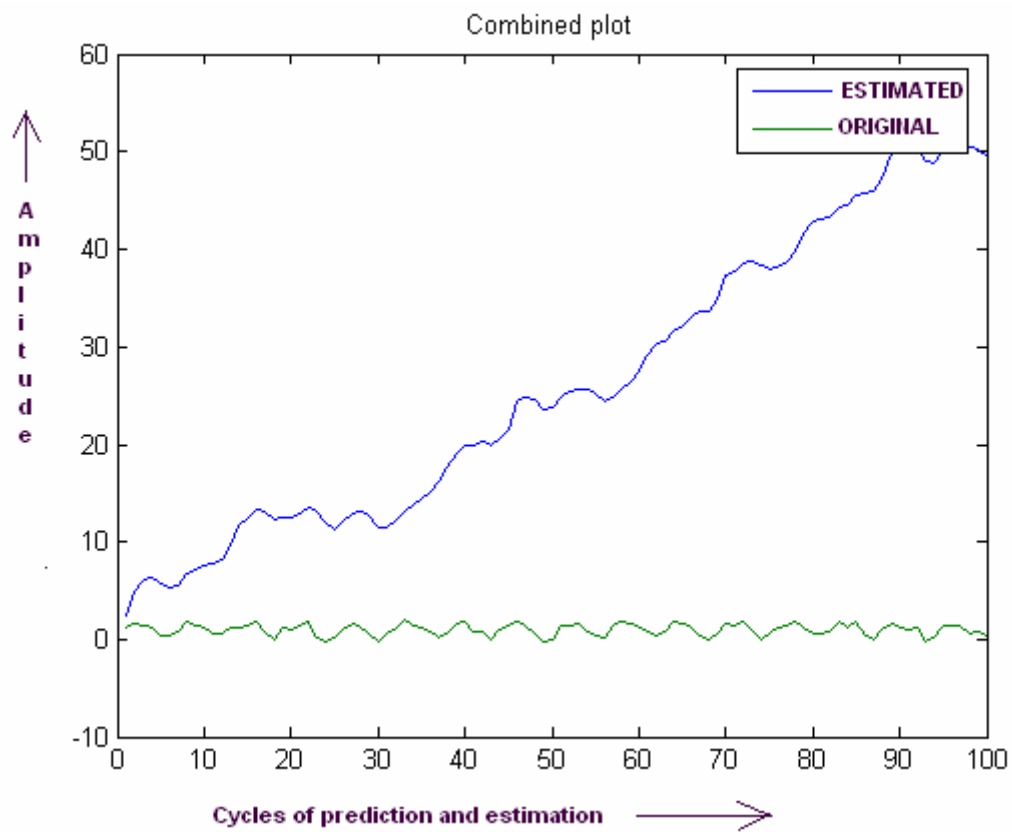


ESTIMATION OF A NOISY SIGNAL OF LOGARITHMIC NATURE



EXTENDED KALMAN FILTER ESTIMATE SIMULATION FOR HIGHER ORDER STATE MODEL

ESTIMATION OF A NOISY SIGNAL OF COSINE NATURE



EXTENDED KALMAN FILTER ESTIMATE SIMULATION FOR HIGHER ORDER STATE MODEL

THE EKF AND ITS FLAWS

Consider the basic state-space estimation framework. Given the noisy observation y_k , a recursive estimation for X_k can be expressed

$$\hat{\mathbf{x}}_k = (\text{prediction of } \mathbf{x}_k) + \mathcal{K}_k \cdot [y_k - (\text{prediction of } y_k)] \quad (8)$$

This recursion provides the optimal minimum mean-squared error (MMSE) Estimate for X_k assuming the prior estimate X_{k-1} and current observation y_k IS Gaussian Random Variables (GRV). We need not assume linearity of the Model. The optimal terms in this recursion are given by

$$\hat{\mathbf{x}}_k^- = E[F(\hat{\mathbf{x}}_{k-1}^-, \mathbf{v}_{k-1})] \quad (9)$$

$$\mathcal{K}_k = \mathbf{P}_{\mathbf{x}_k y_k} \mathbf{P}_{\hat{y}_k \hat{y}_k}^{-1} \quad (10)$$

$$\hat{y}_k^- = E[H(\hat{\mathbf{x}}_k^-, \mathbf{n}_k)], \quad (11)$$

where the optimal prediction of \mathbf{x}_k is written as $\hat{\mathbf{x}}_k^-$, and corresponds to the expectation of a nonlinear function of the random variables $\hat{\mathbf{x}}_{k-1}$ and \mathbf{v}_{k-1} (similar interpretation for the optimal prediction \hat{y}_k^-). The optimal gain term

\mathcal{K}_k is expressed as a function of posterior covariance matrices (WITH $\tilde{y}_k = y_k - \hat{y}_k^-$). These terms also require taking expectations of a NONLINEAR FUNCTION of the prior state estimates. The Kalman filter calculates these quantities exactly in the linear case, and can be viewed as an efficient method for analytically propagating a GRV through linear system DYNAMICS. For nonlinear models, however, the EKF *approximates* the optimal terms as:

$$\hat{\mathbf{x}}_k^- \approx F(\hat{\mathbf{x}}_{k-1}^-, \bar{\mathbf{v}}) \quad (12)$$

$$\mathcal{K}_k \approx \hat{\mathbf{P}}_{\mathbf{x}_k y_k} \hat{\mathbf{P}}_{\hat{y}_k \hat{y}_k}^{-1} \quad (13)$$

$$\hat{y}_k^- \approx H(\hat{\mathbf{x}}_k^-, \bar{\mathbf{n}}), \quad (14)$$

where predictions are approximated as simply the function of the prior *mean* value for estimates (no expectation). The covariance are determined by linearizing the dynamic

EQUATIONS $(\mathbf{x}_{k+1} \approx \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{v}_k, y_k \approx \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{n}_k)$, and then determining the posterior

covariance matrices analytically for the linear system. In other words, in the EKF the state distribution is approximated by a GRV which is then propagated analytically through the "first-order" linearization of the nonlinear system. The EKF can be viewed as providing "first-ORDER" APPROXIMATIONS to the optimal terms. These approximations, however, can introduce large errors in the true posterior mean and covariance of the transformed (Gaussian) random variable, which may lead to sub-optimal performance and sometimes divergence of the filter. It is these "flaws" which will be amended using the UKF.

THE EXTENDED KALMAN FILTER

The linearized Kalman filter we derived is fine as far as it goes but has a problem: we need to know the nominal trajectory \bar{x} ahead of time. For some systems we might know a nominal trajectory ahead of time (for example, an aircraft flight with a predetermined flight plan or a robot arm moving in a manufacturing environment), but for other systems we may have no way of knowing a nominal trajectory.

The idea of the EKF (extended Kalman filter) is to use our estimate of x as the nominal trajectory in the linearized Kalman filter. In other words, we set \bar{x} equal to \hat{x} in the linearized Kalman filter. This is a clever bootstrapping approach to state estimation; we use a nominal trajectory to estimate x and use the estimated value of x as the nominal trajectory. After making these substitutions into the linearized Kalman filter equations and going through some mathematical manipulations, we get the following EKF algorithm:

1. The system equations are given as:

$$\text{State equation: } x_{k+1} = f(x_k, u_k) + w_k$$

(22)

$$\text{Output Equation: } y_k = h(x_k) + v_k$$

(23)

2. At each time step, compute the following derivative matrices, evaluated at the current state estimate:

$$\begin{aligned} A_k &= f'(\hat{x}_k, u_k) \\ C_k &= h'(\hat{x}_k) \end{aligned}$$

(24)

Note that the derivatives are taken with respect to x_k and evaluated at $x_k = \bar{x}_k$.

3. Execute the following Kalman filter equations:

$$\begin{aligned} K_k &= P_k C_k^T (C_k P_k C_k^T + R)^{-1} \\ \hat{x}_{k+1} &= f(\hat{x}_k, u_k) + K_k [y_k - h(\hat{x}_k)] \\ P_{k+1} &= A_k (I - K_k C_k) P_k A_k^T + Q \end{aligned} \quad (25)$$

APPLICATIONS OF EXTENDED KALMAN FILTER :-

MOTOR STATE ESTIMATION

To illustrate the use of the EKF, let's use it to estimate the states of a two-phase permanent magnet synchronous motor. We might want to estimate the states so that we can regulate them with a control algorithm, or we might want to estimate the states because we want to know the position or velocity of the motor for some other reason. Let's suppose that we can measure the motor winding currents, and we want to use the EKF to estimate the rotor position and velocity. The system equations are:

$$\begin{aligned}\dot{I}_a &= \frac{-R}{L} I_a + \frac{\omega \lambda}{L} \sin \theta + \frac{u_a + \Delta u_a}{L} \\ \dot{I}_b &= \frac{-R}{L} I_b + \frac{\omega \lambda}{L} \cos \theta + \frac{u_b + \Delta u_b}{L} \\ \dot{\omega} &= \frac{-3\lambda}{2J} I_a \sin \theta + \frac{3\lambda}{2J} I_b \cos \theta - \frac{F\omega}{J} + \Delta \alpha \\ \dot{\theta} &= \omega \\ y &= \begin{bmatrix} I_a \\ I_b \end{bmatrix} + \begin{bmatrix} v_a \\ v_b \end{bmatrix}\end{aligned}$$

(26)

The variables in these equations are defined as follows:

- ☐ I_a and I_b are the currents in the two motor windings
- ☐ θ and ω are the angular position and velocity of the rotor
- ☐ R and L are the motor winding's resistance and inductance
- ☐ λ is the flux constant of the motor
- ☐ F is the coefficient of viscous friction that acts on the motor shaft and its load
- ☐ J is the moment of inertia of the motor shaft and its load
- ☐ u_a and u_b are the voltages that are applied across the two motor windings
- ☐ Δu_a and Δu_b are noise terms due to errors in u_a and u_b
- ☐ $\Delta \alpha$ is a noise term due to uncertainty in the load torque
- ☐ y is the measurement. We're assuming here that we have measurements of the two winding currents, maybe using sense resistors. The measurements are distorted by measurement noises v_a and v_b which are caused by events such as sense-resistance uncertainty, electrical noise, and quantization errors in our microcontroller.

If we want to apply an EKF to the motor, we need to define the states of the system. You can see the states by looking at the system equations and noting wherever a derivative appears. If a variable is differentiated in the system equations, that quantity is a state. So we see from the motor equations shown in Equation 26 that our system has four states, and the state vector x can be defined as:

$$x = \begin{bmatrix} I_a \\ I_b \\ \omega \\ \theta \end{bmatrix}$$

(27)

The system equation is obtained by "discretizing" the differential equations to obtain the equation in Equation 28 where Δt is the step size that we're using for estimation in our microcontroller or DSP.

$$\begin{aligned} x_{k+1} &= f(x_k, u_k) + w_k \\ &= x_k + \begin{bmatrix} -Rx_k(1)/L + x_k(3)\lambda \sin x_k(4)/L + u_{ak}/L \\ -Rx_k(2)/L + x_k(3)\lambda \cos x_k(4)/L + u_{bk}/L \\ -3\lambda x_k(1)\sin x_k(4)/2J + 3\lambda x_k(2)\cos x_k(4)/2J - Fx_k(3)/J \\ x_k(3) \end{bmatrix} \Delta t + \begin{bmatrix} \Delta u_{ak}/L \\ \Delta u_{bk}/L \\ \Delta \alpha \\ 0 \end{bmatrix} \Delta t \\ y_k &= h(x_k) + v_k \\ &= \begin{bmatrix} x_k(1) \\ x_k(2) \end{bmatrix} + \begin{bmatrix} v_{ak} \\ v_{bk} \end{bmatrix} \end{aligned}$$

(28)

To use an EKF, we need to find the derivatives of $f(x_k, u_k)$ and $h(x_k)$ with respect to x_k because $f(x_k, u_k)$ and $h(x_k)$ are both vector functions of the vector x_k . How can we find the derivative of a vector with respect to another vector? If you already know how to take derivatives, it's really not too difficult to find the derivative of a vector. For example, if we write the vectors f and x as:

$$f = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

(29)

then the derivative of f with respect to x is equal to the 4×4 matrix, shown in Equation 30:

$$f'(x) = \begin{bmatrix} df_1/dx_1 & df_1/dx_2 & df_1/dx_3 & df_1/dx_4 \\ df_2/dx_1 & df_2/dx_2 & df_2/dx_3 & df_2/dx_4 \\ df_3/dx_1 & df_3/dx_2 & df_3/dx_3 & df_3/dx_4 \\ df_4/dx_1 & df_4/dx_2 & df_4/dx_3 & df_4/dx_4 \end{bmatrix}$$

(30)

This matrix can be generalized for any size vectors f and x . With this background, we can now find the derivative matrices as shown in Equation 31:

$$A_k = f'(\hat{x}_k, u_k) = \begin{bmatrix} -R/L & 0 & \dot{\lambda} \sin \hat{x}_k(4)/L & \dot{x}_k(3)\dot{\lambda} \cos \hat{x}_k(4)/L \\ 0 & -R/L & -\dot{\lambda} \cos \hat{x}_k(4)/L & \dot{x}_k(3)\dot{\lambda} \sin \hat{x}_k(4)/L \\ -3\dot{\lambda} \sin \hat{x}_k(4)/2J & 3\dot{\lambda} \cos \hat{x}_k(4)/2J & -F/J & -3\dot{\lambda}[\dot{x}_k(1)\cos \hat{x}_k(4) + \dot{x}_k(2)\sin \hat{x}_k(4)]/2J \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$C_k = h'(\hat{x}_k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

(31)

To simulate the EKF to see how well we can estimate rotor position and velocity. We'll assume that the measurement noise terms, v_{ak} and v_{bk} are zero-mean random variables with standard deviations equal to 0.1 amps. The control inputs (winding voltages) are equal to:

$$u_a(t) = \sin 2\pi t$$

$$u_b(t) = \cos 2\pi t$$

(32)

Equation 32 means that in discrete time, the control inputs are equal to:

$$u_{ak} = \sin 2\pi k\Delta t$$

$$u_{bk} = \cos 2\pi k\Delta t$$

(33)

The voltages that are applied to the winding currents are equal to these values plus u_{ak} and u_{bk} which we'll suppose are zero-mean random variables with standard deviations equal to 0.001 amps. We'll also assume that the noise due to load torque disturbances α_k has a standard deviation of 0.05 rad/sec^2 . Even though our measurements consist only of the winding currents, we can use an EKF to estimate the rotor position and velocity. We see that the rotor position and velocity are estimated quite well.

It's important to know about the limitations of the EKF. For example, Figure 3 was generated using a 1ms sample rate for measurements. If the sample time grows to 2ms, the EKF estimate "blows up;" that is, the EKF becomes unstable. If the measurement noise increases by a factor of 10, the EKF gives position estimates that are essentially worthless. These aren't really limitations of the EKF but rather are limits to the information we can squeeze out of a certain system.

CHAPTER 5

UNSCENTED KALMAN FILTER

UNSCENTED KALMAN FILTER

State Estimation

In order to illustrate the UKF for state-estimation, we provide a new application example corresponding to noisy time-series estimation.

In this example, the UKF is used to estimate an underlying clean time-series corrupted by additive Gaussian white noise. The time-series used is the Mackey-Glass-30 chaotic series. The clean times-series is first modeled as a nonlinear auto regression

$$x_k = f(x_{k-1}, \dots, x_{k-M}, \mathbf{w}) + v_k \quad (19)$$

Where the model f (parameterized by \mathbf{w}) was approximated by training a feed forward neural network on the clean sequence. The residual error after convergence was taken to be the process noise variance.

Next, white Gaussian noise was added to the clean Mackey-Glass series to generate a noisy time-series $y_k = x_k + n_k$. The corresponding state-space representation is given by:

UKF dual estimation

Recall that the dual estimation problem consists of simultaneously estimating the clean state \mathbf{x}_k and the model parameters \mathbf{w} from the noisy data y_k (see Equation 7). As expressed earlier, a number of algorithmic approaches exist for this problem. We present results for the Dual UKF and Joint UKF. Development of a Unscented Smoother for an EM is presented. As in the prior state-estimation example, we utilize a noisy time-series application modeled with neural networks for illustration of the approaches.

In the the *dual extended Kalman filter* 10 a separate state-space representation is used for the signal and the weights. The state-space

Representation for the state \mathbf{x}_k is the same as in Equation 21. In the context of a time-series, the state-space representation for the weights is given by

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \mathbf{u}_k \quad (21)$$

$$y_k = f(\mathbf{x}_{k-1}, \mathbf{w}_k) + v_k + n_k. \quad (22)$$

Where we set the innovations covariance \mathbf{P}_u equal to $\mu \mathbf{P}_w$. Two EKFs can now be run simultaneously for signal and weight estimation. At every time-step, the current estimate of the

weights is used in the signal-filter, and the current estimate of the signal-state is used in the weight-filter. In the new dual UKF algorithm, both state- and weight-estimation are done with the UKF. Note that the state-transition is linear in the weight filter, so the nonlinearity is restricted to the measurement equation.

In the *joint extended Kalman filter*, the signal-state and weight vectors are concatenated into a

$$[\mathbf{x}_k^T \mathbf{w}_k^T]^T$$

single, *joint* state vector: . Estimation is done recursively by writing the state-space equations for the joint state as:

$$\begin{bmatrix} \mathbf{x}_k \\ \mathbf{w}_k \end{bmatrix} = \begin{bmatrix} F(\mathbf{x}_{k-1}, \mathbf{w}_{k-1}) \\ \mathbf{I} \cdot \mathbf{w}_{k-1} \end{bmatrix} + \begin{bmatrix} B \cdot v_k \\ \mathbf{u}_k \end{bmatrix} \quad (23)$$

$$y_k = [1 \quad 0 \quad \dots \quad 0] \begin{bmatrix} \mathbf{x}_k \\ \mathbf{w}_k \end{bmatrix} + n_k, \quad (24)$$

and running an EKF on the joint state-space to produce simultaneous estimates of the states \mathbf{x}_k and \mathbf{w} . SO, our approach will be to use the UKF instead of the EKF to estimate non-linear parameters.

The UKF addresses the approximation issues of the EKF. The state distribution is again represented by a GRV, but is now specified using a minimal set of carefully chosen sample points. These sample points completely capture the true mean and covariance of the GRV, and when propagated through the *true* non-linear system, captures the posterior mean and covariance accurately to the 3rd order (Taylor series expansion) for *any* nonlinearity. To elaborate on this, we start by first explaining the *unscented transformation*. The unscented transformation (UT) is a method for calculating the statistics of a random variable which undergoes a nonlinear transformation. Consider propagating a random variable \mathbf{x} (dimension L) through a nonlinear

function, $\mathbf{y} = g(\mathbf{x})$. Assume \mathbf{x} has mean $\bar{\mathbf{x}}$ and covariance \mathbf{P}_x . To calculate the statistics of \mathbf{y} , we form a matrix \mathcal{X} of $2L+1$ sigma vectors \mathcal{X}_i (with corresponding weights W_i), according to the following:

$$\mathcal{X}_0 = \bar{\mathbf{x}} \quad (15)$$

$$\mathcal{X}_i = \bar{\mathbf{x}} + \left(\sqrt{(L+\lambda)\mathbf{P}_x} \right)_i \quad i = 1, \dots, L$$

$$\mathcal{X}_i = \bar{\mathbf{x}} - \left(\sqrt{(L+\lambda)\mathbf{P}_x} \right)_{i-L} \quad i = L+1, \dots, 2L$$

$$W_0^{(m)} = \lambda / (L + \lambda)$$

$$W_0^{(c)} = \lambda / (L + \lambda) + (1 - \alpha^2 + \beta)$$

$$W_i^{(m)} = W_i^{(c)} = 1 / \{2(L + \lambda)\} \quad i = 1, \dots, 2L$$

$$\lambda = \alpha^2(L + \kappa) - L$$

Where α is a scaling parameter. α Determines the spread of the sigma points around $\bar{\mathbf{x}}$ and is usually set to a small positive value (e.g., 1e-3). κ is a secondary scaling

parameter which is usually set to 0, and β is used to incorporate prior knowledge of the

distribution of \mathbf{x} (for Gaussian distributions, $\beta = 2$ is optimal). $(\sqrt{(L + \lambda)\mathbf{P}_x})_i$ is the i th row of the matrix square root. These sigma vectors are propagated through the nonlinear function,

$$\mathcal{Y}_i = g(\mathcal{X}_i) \quad i = 0, \dots, 2L, \quad (16)$$

And the mean and covariance for \mathbf{y} are approximated using a weighted sample mean and covariance of the posterior sigma points,

$$\bar{\mathbf{y}} \approx \sum_{i=0}^{2L} W_i^{(m)} \mathcal{Y}_i \quad (17)$$

$$\mathbf{P}_y \approx \sum_{i=0}^{2L} W_i^{(c)} \{\mathcal{Y}_i - \bar{\mathbf{y}}\} \{\mathcal{Y}_i - \bar{\mathbf{y}}\}^T \quad (18)$$

This method differs substantially from general "sampling" methods which require orders of magnitude more sample points in an attempt to propagate an accurate (possibly non-Gaussian) distribution of the state. The deceptively simple approach taken with the UT results in approximations that is accurate to the third order for Gaussian inputs for all nonlinearities. For non-Gaussian inputs, approximations are accurate to at least the second-order, with the accuracy of third and higher order moments determined by the choice of α and β .

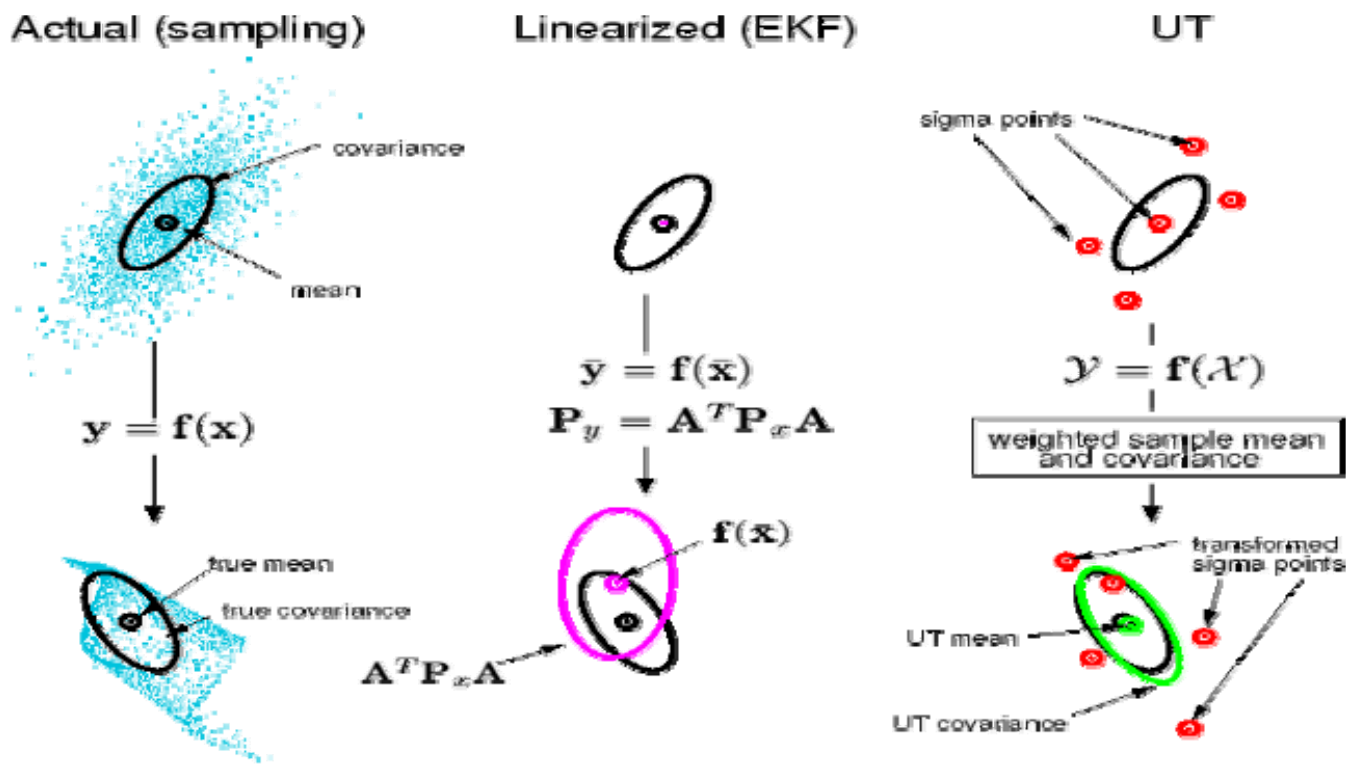


Figure 1.2: Example of the UT for mean and covariance propagation. a) Actual, b) first-order linearization (EKF), c) UT.

The left plot shows the true mean and covariance propagation using Monte-Carlo sampling; the center plots show the results using a linearization approach as would be done in the EKF; the right plots show the performance of the UT (note only 5 sigma points are required). The superior performance of the UT is clear. The Unscented Kalman Filter (UKF) is a straightforward extension of the UT to the recursive estimation in Equation 8, where the state RV

$$\mathbf{x}_k^a = [\mathbf{x}_k^T \mathbf{v}_k^T \mathbf{n}_k^T]^T$$

is redefined as the concatenation of the original state and noise variables:

The UT sigma point selection scheme (Equation 15) is applied to this new augmented state RV

$$\mathbf{x}_k^a$$

to calculate the corresponding sigma matrix, \mathbf{P}_k^a . The UKF equations are given in

Algorithm requires no explicit calculations of Jacobians or Hessians are necessary to implement this algorithm. Furthermore, the overall number of computations is the same order as the EKF.

UKF PROCESS

Initialize with:

$$\begin{aligned}\hat{\mathbf{x}}_0 &= E[\mathbf{x}_0] \\ \mathbf{P}_0 &= E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T] \\ \hat{\mathbf{x}}_0^a &= E[\mathbf{x}^a] = [\hat{\mathbf{x}}_0^T \mathbf{0} \mathbf{0}]^T \\ \mathbf{P}_0^a &= E[(\mathbf{x}_0^a - \hat{\mathbf{x}}_0^a)(\mathbf{x}_0^a - \hat{\mathbf{x}}_0^a)^T] = \begin{bmatrix} \mathbf{P}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_v & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_n \end{bmatrix}\end{aligned}$$

For $k \in \{1, \dots, \infty\}$,

Calculate sigma points:

$$\mathcal{X}_{k-1}^a = [\hat{\mathbf{x}}_{k-1}^a \quad \hat{\mathbf{x}}_{k-1}^a \pm \sqrt{(L + \lambda) \mathbf{P}_{k-1}^a}]$$

Time update:

$$\begin{aligned}\mathcal{X}_{k|k-1}^x &= \mathbf{F}[\mathcal{X}_{k-1}^x, \mathcal{X}_{k-1}^v] \\ \hat{\mathbf{x}}_k^- &= \sum_{i=0}^{2L} W_i^{(m)} \mathcal{X}_{i,k|k-1}^x \\ \mathbf{P}_k^- &= \sum_{i=0}^{2L} W_i^{(c)} [\mathcal{X}_{i,k|k-1}^x - \hat{\mathbf{x}}_k^-][\mathcal{X}_{i,k|k-1}^x - \hat{\mathbf{x}}_k^-]^T \\ \mathcal{Y}_{k|k-1} &= \mathbf{H}[\mathcal{X}_{k|k-1}^x, \mathcal{X}_{k-1}^n] \\ \hat{\mathbf{y}}_k^- &= \sum_{i=0}^{2L} W_i^{(m)} \mathcal{Y}_{i,k|k-1}\end{aligned}$$

Measurement update equations:

$$\begin{aligned}\mathbf{P}_{\hat{\mathbf{y}}_k \hat{\mathbf{y}}_k} &= \sum_{i=0}^{2L} W_i^{(c)} [\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_k^-][\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_k^-]^T \\ \mathbf{P}_{\mathbf{x}_k \mathbf{y}_k} &= \sum_{i=0}^{2L} W_i^{(c)} [\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}_k^-][\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_k^-]^T \\ \mathbf{K} &= \mathbf{P}_{\mathbf{x}_k \mathbf{y}_k} \mathbf{P}_{\hat{\mathbf{y}}_k \hat{\mathbf{y}}_k}^{-1} \\ \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_k^- + \mathbf{K}(\mathbf{y}_k - \hat{\mathbf{y}}_k^-) \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K} \mathbf{P}_{\hat{\mathbf{y}}_k \hat{\mathbf{y}}_k} \mathbf{K}^T\end{aligned}$$

where, $\mathbf{x}^a = [\mathbf{x}^T \mathbf{v}^T \mathbf{n}^T]^T$, $\mathcal{X}^a = [(\mathcal{X}^x)^T (\mathcal{X}^v)^T (\mathcal{X}^n)^T]^T$, λ =composite scaling parameter, L =dimension of augmented state, \mathbf{P}_v =process noise cov., \mathbf{P}_n =measurement noise cov., W_i =weights as calculated in Eqn. 15.

UNSCENTED KALMAN FILTER ALGORITHM

UNSCENTED KALMAN FILTER

The extended Kalman filter gives particularly poor performance on highly non-linear functions because only the mean is propagated through the non-linearity. The unscented Kalman filter (UKF) uses a deterministic sampling technique to pick a minimal set of sample points (called sigma points) around the mean. These sigma points are then propagated through the non-linear functions and the covariance of the estimate is then recovered. The result is a filter which more accurately captures the true mean and covariance. (This can be verified using Monte Carlo sampling or through a Taylor series expansion of the posterior statistics.) In addition, this technique removes the requirement to analytically calculate Jacobians, which for complex functions can be a difficult task in itself.

PREDICT :

As with the EKF, the UKF prediction can be used independently from the UKF update, in combination with a linear (or indeed EKF) update, or vice versa.

The estimated state and covariance are augmented with the mean and covariance of the process noise.

$$\mathbf{x}_{k-1|k-1}^a = [\hat{\mathbf{x}}_{k-1|k-1}^T \quad E[\mathbf{w}_k^T]]^T$$
$$\mathbf{P}_{k-1|k-1}^a = \begin{bmatrix} \mathbf{P}_{k-1|k-1} & 0 \\ 0 & \mathbf{Q}_k \end{bmatrix}$$

A set of $2L+1$ sigma points is derived from the augmented state and covariance where L is the dimension of the augmented state.

$$\chi_{k-1|k-1}^0 = \mathbf{x}_{k-1|k-1}^a$$
$$\chi_{k-1|k-1}^i = \mathbf{x}_{k-1|k-1}^a + \left(\sqrt{(L + \lambda) \mathbf{P}_{k-1|k-1}^a} \right)_i \quad i = 1..L$$
$$\chi_{k-1|k-1}^i = \mathbf{x}_{k-1|k-1}^a - \left(\sqrt{(L + \lambda) \mathbf{P}_{k-1|k-1}^a} \right)_{i-L} \quad i = L + 1, \dots, 2L$$

The sigma points are propagated through the transition function f .

$$\chi_{k|k-1}^i = f(\chi_{k-1|k-1}^i) \quad i = 0..2L$$

The weighted sigma points are recombined to produce the predicted state and covariance.

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{i=0}^{2L} W_s^i \chi_{k|k-1}^i$$

$$\mathbf{P}_{k|k-1} = \sum_{i=0}^{2L} W_c^i [\chi_{k|k-1}^i - \hat{\mathbf{x}}_{k|k-1}][\chi_{k|k-1}^i - \hat{\mathbf{x}}_{k|k-1}]^T$$

Where the weights for the state and covariance are given are:

$$\begin{aligned} W_s^0 &= \frac{\lambda}{L + \lambda} \\ W_c^0 &= \frac{\lambda}{L + \lambda} + (1 - \alpha^2 + \beta) \\ W_s^i &= W_c^i = \frac{1}{2(L + \lambda)} \\ \lambda &= \alpha^2(L + \kappa) - L \end{aligned}$$

Typical values for α , β , and κ are 10^{-3} , 2 and 0 respectively. (These values should suffice for most purposes.)

UPDATE :

The predicted state and covariance are augmented as before, except now with the mean and covariance of the measurement noise.

$$\begin{aligned} \mathbf{x}_{k|k-1}^a &= [\hat{\mathbf{x}}_{k|k-1}^T \quad E[\mathbf{v}_k^T]]^T \\ \mathbf{P}_{k|k-1}^a &= \begin{bmatrix} \mathbf{P}_{k|k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_k \end{bmatrix} \end{aligned}$$

As before, a set of $2L + 1$ sigma points is derived from the augmented state and covariance where L is the dimension of the augmented state.

$$\begin{aligned} \chi_{k|k-1}^0 &= \mathbf{x}_{k|k-1}^a \\ \chi_{k|k-1}^i &= \mathbf{x}_{k|k-1}^a + \left(\sqrt{(L + \lambda) \mathbf{P}_{k|k-1}^a} \right)_i \quad i = 1..L \\ \chi_{k|k-1}^i &= \mathbf{x}_{k|k-1}^a - \left(\sqrt{(L + \lambda) \mathbf{P}_{k|k-1}^a} \right)_{i-L} \quad i = L + 1, \dots, 2L \end{aligned}$$

ALTERNATIVELY IF THE UKF PREDICTION HAS BEEN USED THE SIGMA POINTS THEMSELVES CAN BE AUGMENTED ALONG THE FOLLOWING LINES

$$\chi_{k|k-1} := [\chi_{k|k-1} \quad E[\mathbf{v}_k^T]]^T \pm \sqrt{(L + \lambda) \mathbf{R}_k^a}$$

WHERE

$$\mathbf{R}_k^a = \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{R}_k \end{bmatrix}$$

THE SIGMA POINTS ARE PROJECTED THROUGH THE OBSERVATION FUNCTION H.

$$\gamma_k^i = h(\chi_{k|k-1}^i) \quad i = 0..2L$$

THE WEIGHTED SIGMA POINTS ARE RECOMBINED TO PRODUCE THE PREDICTED MEASUREMENT AND PREDICTED MEASUREMENT COVARIANCE.

$$\hat{\mathbf{z}}_k = \sum_{i=0}^{2L} W_s^i \gamma_k^i$$

$$\mathbf{P}_{z_k z_k} = \sum_{i=0}^{2L} W_c^i [\gamma_k^i - \hat{\mathbf{z}}_k][\gamma_k^i - \hat{\mathbf{z}}_k]^T$$

THE STATE-MEASUREMENT CROSS-CORRELATION MATRIX,

$$\mathbf{P}_{x_k z_k} = \sum_{i=0}^{2L} W_c^i [\chi_{k|k-1}^i - \hat{\mathbf{x}}_{k|k-1}][\gamma_k^i - \hat{\mathbf{z}}_k]^T$$

IS USED TO COMPUTE THE UKF KALMAN GAIN.

$$\mathbf{K}_k = \mathbf{P}_{x_k z_k} \mathbf{P}_{z_k z_k}^{-1}$$

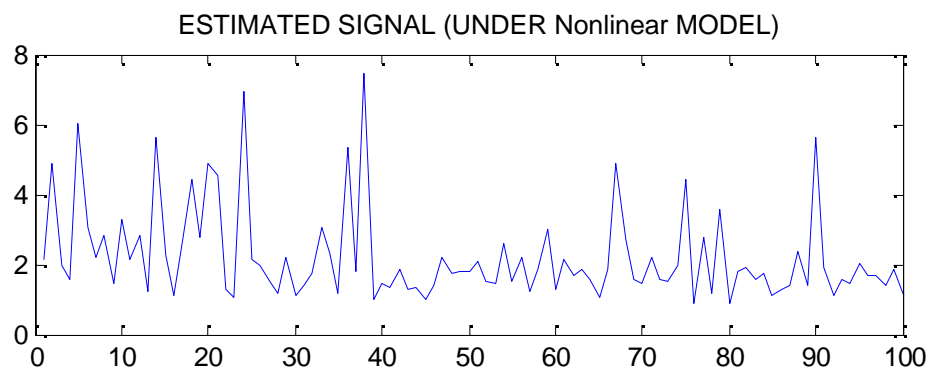
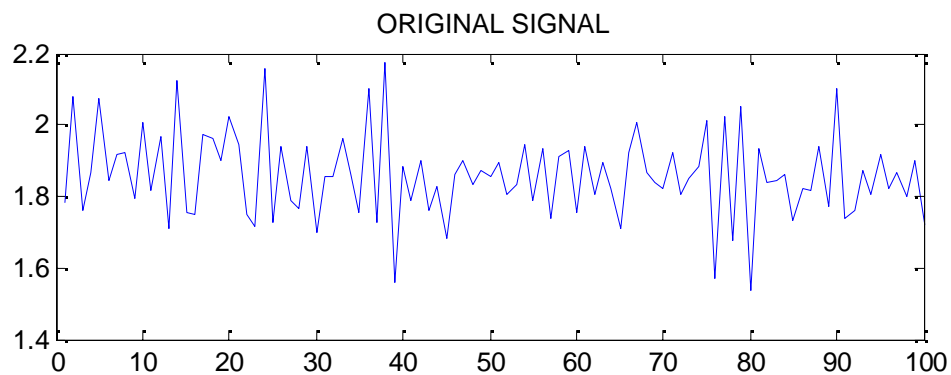
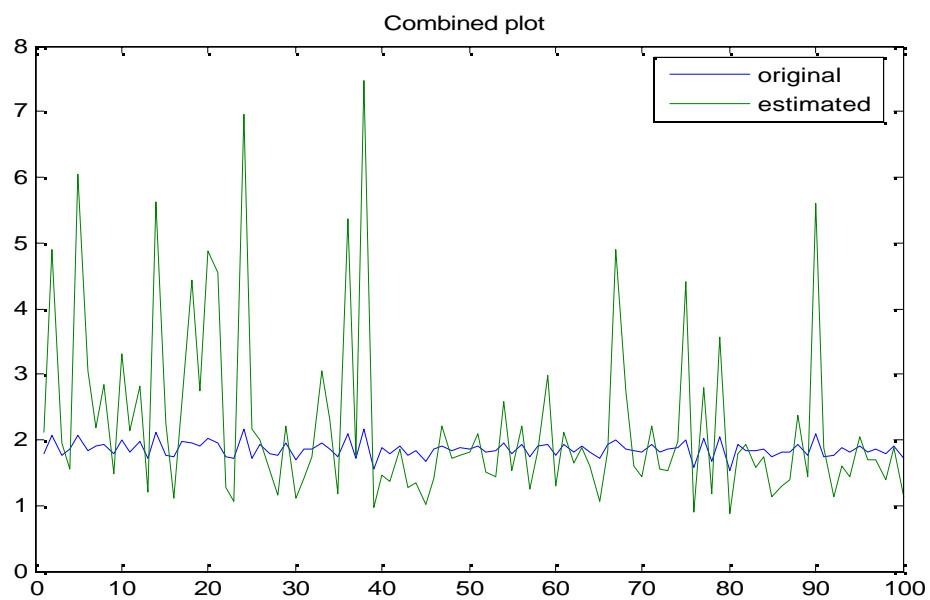
AS WITH THE KALMAN FILTER, THE UPDATED STATE IS THE PREDICTED STATE PLUS THE INNOVATION WEIGHTED BY THE KALMAN GAIN,

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_k)$$

AND THE UPDATED COVARIANCE IS THE PREDICTED COVARIANCE, MINUS THE PREDICTED MEASUREMENT COVARIANCE, WEIGHTED BY THE KALMAN GAIN.

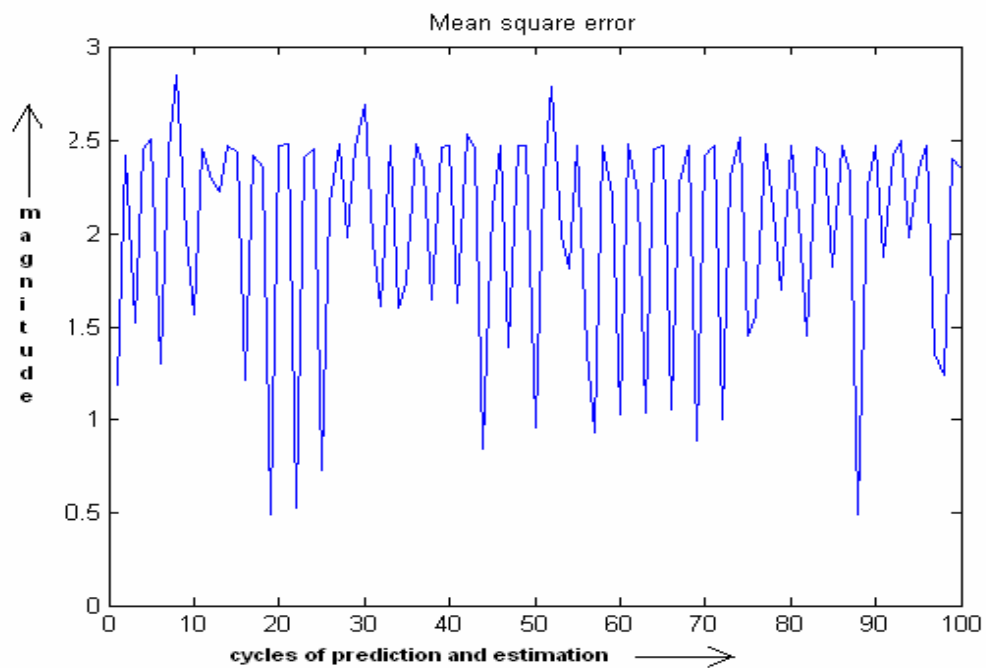
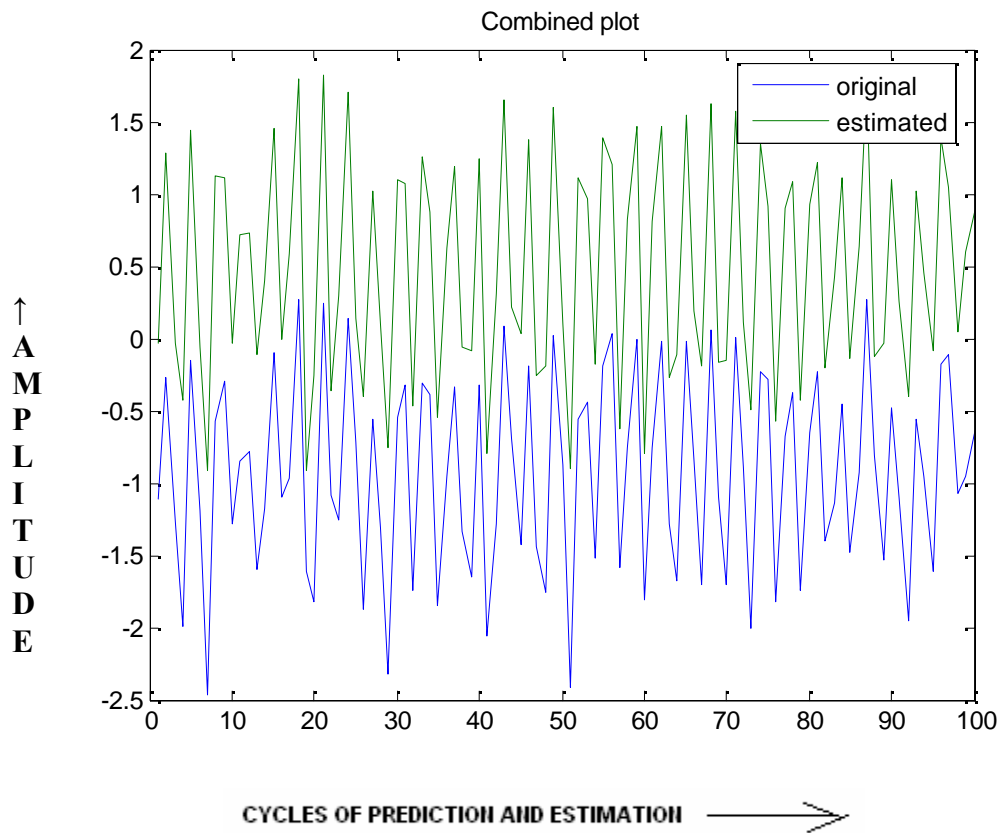
$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{P}_{z_k z_k} \mathbf{K}_k^T$$

UNSCENTED KALMAN FILTER SIMULATION OUTPUT



UNSCENTED KALMAN FILTER ESTIMATE SIMULATION

UNSCETED KALMAN FILTER ESTIMATION IN HIGH NON-LINEARITIES



**UNSCENTED KALMAN FILTER ESTIMATE SIMULATION FOR
HIGHER ORDER STATE MODEL**

UKF FILTER IMPLEMENTATION TO ESTIMATE POWER FREQUENCY:-

INTRODUCTION

The problem of propagating Gaussian random variables through a nonlinear function can also be approached using another technique, namely the *unscented transform*. Instead of linearizing the functions, this transform uses a set of points, and propagates them through the actual nonlinear function, eliminating linearization altogether. The points are chosen such that their mean, covariance, and possibly also higher order moments, match the Gaussian random variable. Mean and covariance can be recalculated from the propagated points, yielding more accurate results compared to ordinary function linearization. The underlying idea is also to approximate the probability distribution instead of the function. This strategy typically does both decrease the computational complexity, while at the same time increasing estimate accuracy, yielding faster, more accurate results.

The underlying method of unscented transform was first proposed by Uhlmann et al., where they laid out the framework for representing a Gaussian random variable in N dimensions using $2N + 1$ samples, called *sigma points*. They utilized the matrix square root and covariance definitions to select these points in such a way that they had the same covariance as the Gaussian they approximated. Skewness was avoided by selecting the points in a symmetric way, such that any approximation error would only originate from the fourth and higher moments. Usage of the unscented transform in Kalman filtering was then presented by Julier, where he introduced the *Unscented Kalman filter* (UKF), which approximates the state estimate using sigma points. A limitation associated with the unscented Kalman filter is that it has a lower bound on the *safe spread* of the sigma points, meaning the distance between the points in state space. Sigma point spreads below this bound are not guaranteed to yield positive semi definite correlation matrices. This distance also increases with the dimension of the state space, a limitation that may cause problems in highly nonlinear models, since high sigma point spread may result in sampling of non-local features. The technique presented here is therefore based on the *scaled unscented transform*, which provides an additional tuning parameter, compared to the original Unscented transform. This parameter is used to arbitrarily control the spread of the sigma-points, while at the same time guaranteeing positive semi definite covariance matrices. Even models of high dimensionality can then keep a tight sigma point spread to avoid non local effects.

AUGMENTED STATE

The unscented transform approach also has another advantage, namely that noise can be treated in a nonlinear fashion to account for non-Gaussian or nonadditive noises. The strategy for doing so involves propagation of noise through the functions by first augmenting the state vector to also include noise sources, a technique first introduced by Julier, and later refined more in depth by Merwe. Sigma point samples are then selected from the augmented state, \mathbf{x}_a , which also includes noise values. The net result is that any nonlinear effects of process and measurement noise are captured with the same accuracy as the rest of the state, which in turn increases accuracy for non-additive noise sources.

FILTER FORMULATION

The filter starts by augmenting the state vector to L dimensions, where L is the sum of dimensions in the original state-vector, model noise and measurement noise. The covariance matrix is similarly augmented

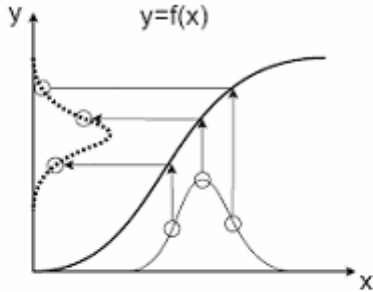


Figure 7.1 Illustration of how the unscented Kalman filter propagates sigma-points from a Gaussian distribution through a nonlinear function, and recreates a Gaussian distribution, by calculating the mean and covariance of the results

to a $2L+1$ matrix. Together this forms the augmented state estimate vector \mathbf{x}^a and covariance matrix P^a :

$$\mathbf{x}_{k-1}^a = \begin{bmatrix} \mathbf{x}_{k-1} \\ 0_w \\ 0_v \end{bmatrix}$$

$$P_{k-1}^a = E\{(\mathbf{x}_{k-1}^a - \hat{\mathbf{x}}_{k-1}^a)(\mathbf{x}_{k-1}^a - \hat{\mathbf{x}}_{k-1}^a)^T\}$$

$$= \begin{bmatrix} P_{k-1} & 0 & 0 \\ 0 & Q_{k-1} & 0 \\ 0 & 0 & R_{k-1} \end{bmatrix}$$

The next step consists of creating $2L+1$ sigma-points in such a way that they together captures the full mean and covariance of the augmented state. The χ^a matrix is chosen to contain these points, and its columns are calculated as follows:

$$\chi_{0,k-1}^a = \mathbf{x}_{k-1}^a \quad i = 0$$

$$\chi_{i,k-1}^a = \mathbf{x}_{k-1}^a + (\alpha \sqrt{L P_{k-1}^a})_i, \quad i = 1, \dots, L$$

$$\chi_{i,k-1}^a = \mathbf{x}_{k-1}^a - (\alpha \sqrt{L P_{k-1}^a})_{i-L}, \quad i = L+1, \dots, 2L$$

Where subscript i mean the i -th column of the square root of the covariance matrix². The parameter, in the interval $0-1$, determines sigma-point spread. This parameter is typically quite low, often around 0.001, to avoid non-local effects. The resulting χ_{k-1}^a matrix can now be decomposed vertically into the x_{k-1} rows, which contains the state; the w_{k-1} rows, which contains sampled process noise and the v_{k-1} rows, which contains sampled measurement noise. Each sigma-point is also assigned a weight. These weights are derived by comparing the moments of the sigma-points with a Taylor series expansion of the models while assuming a

Gaussian distribution, as derived. The resulting weights for mean (m) and covariance (c) estimates then become:

$$\begin{aligned} w_0^{(m)} &= 1 - \frac{1}{\alpha^2} & i &= 0 \\ w_0^{(c)} &= 4 - \frac{1}{\alpha^2} - \alpha^2 & i &= 0 \\ w_i^{(m)} &= w_i^{(c)} = \frac{1}{2\alpha^2 L} & i &= 1, \dots, 2L \end{aligned}$$

The filter then predicts next state by propagating the sigma-points through the state and measurement models, and then calculating weighted averages and covariance matrices of the results:

$$\begin{aligned} \chi_{k|k-1}^x &= f(\chi_{k-1}^x, \mathbf{u}_k, \chi_{k-1}^w) \\ \hat{\mathbf{x}}_{k|k-1} &= \sum_{i=0}^{2L} w_i^{(m)} \chi_{k|k-1}^x \\ P_{k|k-1} &= \sum_{i=0}^{2L} w_i^{(c)} [\chi_{k|k-1}^x - \hat{\mathbf{x}}_{k|k-1}] [\chi_{k|k-1}^x - \hat{\mathbf{x}}_{k|k-1}]^T \\ Z_{k|k-1} &= h(\chi_{k|k-1}^x, \chi_{k-1}^y) \\ \hat{\mathbf{z}}_{k|k-1} &= \sum_{i=0}^{2L} w_i^{(m)} Z_{i,k|k-1} \end{aligned}$$

The predictions are then updated with new measurements by first calculating the measurement covariance and state-measurement cross correlation matrices, which are then used to determine the Kalman gain:

$$\begin{aligned} P_{zz} &= \sum_{i=0}^{2L} w_i^{(c)} [Z_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1}] [Z_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1}]^T \\ P_{xz} &= \sum_{i=0}^{2L} w_i^{(c)} [\chi_{i,k|k-1}^x - \hat{\mathbf{x}}_{k|k-1}] [Z_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1}]^T \\ K_k &= P_{xz} P_{zz}^{-1} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + K_k (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}) \\ P_{k|k} &= P_{k|k-1} - K_k P_{zz} K_k^T \end{aligned}$$

Experimental results indicate that Unscented Kalman filters yield results comparable to a third order Taylor series expansion of the state-model, while Extended Kalman filters of course only are accurate to a first order linearization.

STEPS FOR UNSCENTED KALMAN FILTERING :-

1. The set of sigma points are created to the augmented system

2. The transformed set is given by instantiating each point through the process model,

$$\mathcal{X}_i(k+1|k) = \mathbf{f}[\mathcal{X}_i^a(k|k), \mathbf{u}(k), k].$$

3. The predicted mean is computed as

$$\hat{\mathbf{x}}(k+1|k) = \sum_{i=0}^{2n^a} W_i \mathcal{X}_i^a(k+1|k).$$

4. And the predicted covariance is computed as

$$\mathbf{P}(k+1|k) = \sum_{i=0}^{2n^a} W_i \{\mathcal{X}_i(k+1|k) - \hat{\mathbf{x}}(k+1|k)\} \{\mathcal{X}_i(k+1|k) - \hat{\mathbf{x}}(k+1|k)\}^T$$

5. Instantiate each of the prediction points through the observation model,

$$\mathcal{Z}_i(k+1|k) = \mathbf{h}[\mathcal{X}_i(k+1|k), \mathbf{u}(k), k]$$

6. The predicted observation is calculated by

$$\hat{\mathbf{z}}(k+1|k) = \sum_{i=1}^{2n^a} W_i \mathcal{Z}_i(k+1|k).$$

7. Since the observation noise is additive and independent, the innovation covariance is

$$\mathbf{P}_{vv}(k+1|k) = \mathbf{R}(k+1) + \sum_{i=0}^{2n^a} W_i \{\mathcal{Z}_i(k|k-1) - \hat{\mathbf{z}}(k+1|k)\} \{\mathcal{Z}_i(k|k-1) - \hat{\mathbf{z}}(k+1|k)\}^T$$

8. Finally the cross correlation matrix is determined by

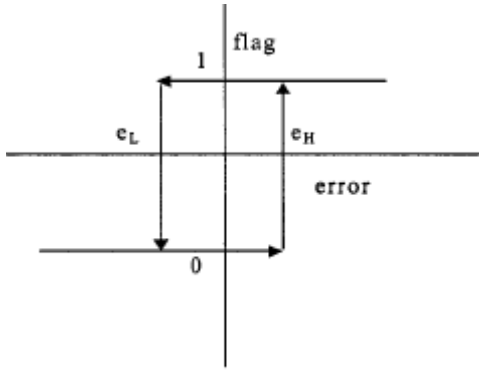
$$\mathbf{P}_{xz}(k+1|k) = \sum_{i=0}^{2n^a} W_i \{\mathcal{X}_i(k|k-1) - \hat{\mathbf{x}}(k+1|k)\} \{\mathcal{Z}_i(k|k-1) - \hat{\mathbf{z}}(k+1|k)\}^T$$

CHAPTER 6

EXTENDED COMPLEX KALMAN FILTERS

EXTENDED COMPLEX KALMAN FILTERS

. The amplitude and phase estimation of a digitized signal has been an important area of research for the past several years, and the methods have been almost standardized for signals with known frequencies. However, if the frequency is not known *a priori* it becomes a formidable task to accurately measure the amplitude and phase. It has been a perpetual problem even now to correctly estimate the frequency of an incoming signal from its sampled values under high noise conditions. The reason for this is the association of severe nonlinearity in the modeling process. On the other hand the philosophy of protection for power systems has been undergoing a change. The modern protective relays employ sophisticated signal processing algorithms to accurately estimate the system conditions from the measured voltage and current signals. The pitfalls in these schemes stem from the frequency measurement algorithm which, under transient and abnormal conditions, cannot sense the correct value. Many algorithms have been reported in the literature for measurement and estimation of frequency.[1]–[4]. A comparative study among four different trackers has been outlined in [5]. The performances of 1) an adaptive notch filter, 2) a multiple frequency tracker, 3) an adaptive IIR filter, and 4)



The hysteresis band for resetting the covariance matrix.

a hyperstable adaptive line enhancer have been presented. The complex Kalman filtering has been used in [6] to estimate the frequency of the signals corrupted with white noise. The present paper is based on the fundamental work carried out in [6]. However, in practice the real and imaginary signals can not be obtained simultaneously. Therefore, keeping in view the measurement of power system signals, the signal model has been modified. The output equation is rewritten to calculate the real signal.

II. SIGNAL MODEL

The nonlinear state space description of the power system signal can have various forms. Numerous linear as well as nonlinear models have been proposed to estimate the amplitude, phase and frequency of a single sinusoid. This signal can also be represented by a complex model. With availability of floating point DSP processors the computation in the complex domain is no longer a difficult task. Moreover, the complex representation is much simpler and direct as far as the frequency measurement is concerned. Once the signal model is established the extended Kalman filter theory [7] can be applied to identify the filter equations and computational steps.

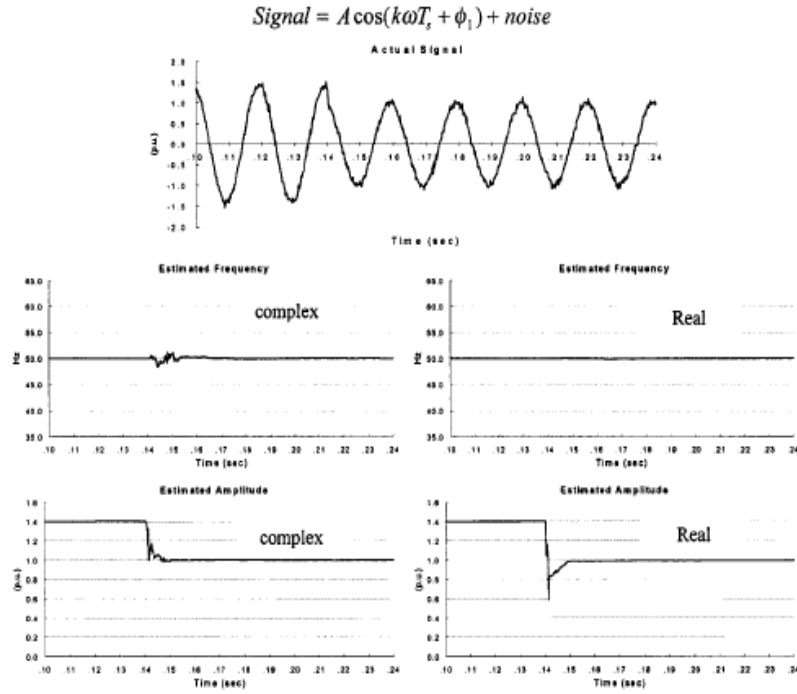
Let an observation signal y_k at time t_k be a sum z_k of M sinusoids with additive noise v_k

$$y_k = z_k + v_k \quad k = 1, 2, 3, \dots, N \quad (1)$$

where

$$z_k = \sum_{i=1}^M a_i \sin(\omega_i t_k + \phi_i) \quad \omega_i = 2\pi f_i, t_k = k\Delta t \quad (2)$$

in which a_i , f_i , and ϕ_i are amplitude, frequency and the phase of the i th sinusoid, respectively. The observation noise v_k is a Gaussian white noise with zero-mean and variance σ_v^2 .



In the case of power system signals, the percentage of frequency components other than fundamental is low. Therefore these harmonic components need not be considered in the model.

For such systems, (2) reduces to a single sinusoid

For such systems, (2) reduces to a single sinusoid

$$z_k = a_1 \sin(k\omega_1 T_s + \phi_1) \quad (3)$$

where T_s = sampling time, and ω_1 = the fundamental angular frequency

The observation signal can be represented in an autoregressive complex form as follows:

$$\begin{bmatrix} \alpha \\ u_{k+1} \\ u_{k+1}^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \frac{1}{\alpha} \end{bmatrix} \begin{bmatrix} \alpha \\ u_k \\ u_k^* \end{bmatrix} \quad (4)$$

$$y_k = [0 \quad 0.5 \quad -0.5] \begin{bmatrix} \alpha \\ u_k \\ u_k^* \end{bmatrix} + v_k \quad (5)$$

where

$$\alpha = \exp(j\omega_1 T_s) \quad (6)$$

$$u_k = a_1 \exp(jk\omega_1 T_s + j\phi_1) \quad (7)$$

$$u_k^* = a_1 \exp(-jk\omega_1 T_s - j\phi_1) \quad (8)$$

v_k = measurement noise.

The above nonlinear process can be represented as

$$x_{k+1} = f(x_k) \quad (9)$$

$$y_{k+1} = Hx_k + v_k \quad (10)$$

where

$$x(k) = [\alpha \quad u_k \quad u_k^*]^T \quad (11)$$

$$f(x_k) = \begin{bmatrix} \alpha & \alpha u_k & \frac{u_k^*}{\alpha} \end{bmatrix}^T \quad (12)$$

$$H = \begin{bmatrix} 0 & 0.5 & -0.5 \end{bmatrix}. \quad (13)$$

Linearizing the above system and applying the extended complex Kalman filter to the first-order system, a nonlinear recursive filter for estimating a single complex sinusoid and its frequency in white noise is obtained as follows [6]:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - H\hat{x}_{k|k-1}) \quad (14)$$

$$\hat{x}_{k+1|k} = f(\hat{x}_{k|k}) \quad (15)$$

$$K_k = \hat{P}_{k|k-1} H^{*T} [H \hat{P}_{k|k-1} H^{*T} + 1]^{-1} \quad (16)$$

$$\hat{P}_{k|k} = \hat{P}_{k|k-1} - K_k H \hat{P}_{k|k-1} \quad (17)$$

$$\hat{P}_{k|k+1} = F_k \hat{P}_{k|k} F_k^{*T} \quad (18)$$

where

$$F_k = \left. \frac{\partial f(x_k)}{\partial x_k} \right|_{x_k = \hat{x}_{k|k}} = \begin{bmatrix} 1 & 0 & 0 \\ \hat{x}_{k|k}(2) & \hat{x}_{k|k}(1) & 0 \\ \frac{-\hat{x}_{k|k}(3)}{\hat{x}_{k|k}^2(1)} & 0 & \frac{1}{\hat{x}_{k|k}(1)} \end{bmatrix}$$

$$\hat{P}_{k|k} = \hat{\Sigma}_{k|k} / \sigma_v^2 \quad \hat{P}_{k+1|k} = \hat{\Sigma}_{k+1|k} / \sigma_v^2$$

$$\hat{\Sigma}_{k|k} = E\{(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^{*T}\}$$

$$\hat{\Sigma}_{k+1|k} = E\{(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^{*T}\}$$

$$\hat{x}_{1|0} = \bar{x}_1, \quad \hat{P}_{1|0} = \hat{\Sigma}_{1|0} / \sigma_v^2$$

$$\hat{\Sigma}_{1|0} = \Sigma_{x1}, \quad \Sigma_{x1} = E\{(x_1 - \bar{x}_1)(x_1 - \bar{x}_1)^{*T}\}$$

CONCLUSION

A nonlinear filter based on the complex Kalman filter has been proposed. The hysteresis method has been suggested for resetting the covariance matrix, which enables fast tracking of the frequency. The proposed filter offers superior performances in all cases. In addition the filter involves less computation which makes it attractive for real-time implementation. The issues such as stability of the algorithm and performance under other disturbances are currently under investigation.

CHAPTER 7

COMPARISION BETWEEN DIFFERENT TYPES OF FILTERS

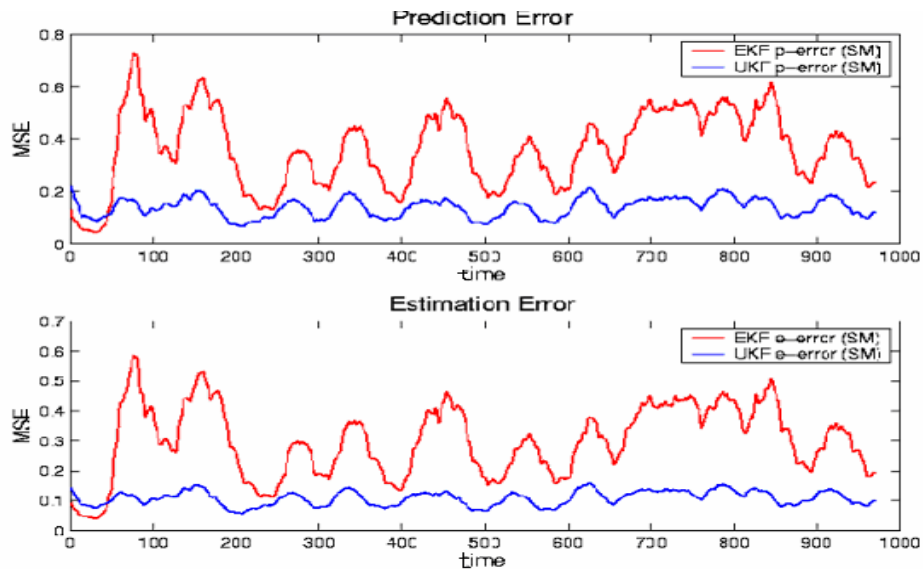
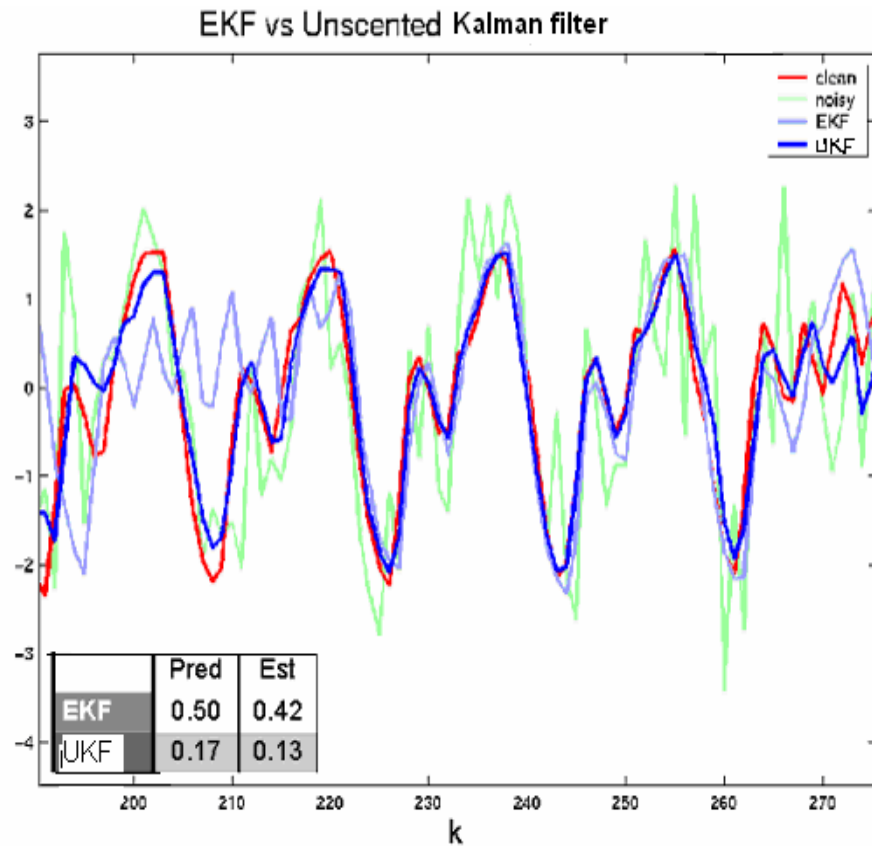
COMPARSION BETWEEN **THE DIFFERENT TYPES OF FILTERS**

The Kalman filter (KF) is an optimal linear estimator when the process noise and the measurement noise can be modeled by white Gaussian noise. The KF only utilizes the first two moments of the state (mean and covariance) in its update rule. It is regarded as an adaptive low-pass infinite impulse response digital filter with cut off frequency depending on ratio between process and measurement noise as well as estimate covariance.

In situations when the problems are nonlinear or the noise that distorts the signals is non-Gaussian, the Kalman filters provide a solution that may be far from optimal. Nonlinear problems can be solved with the extended Kalman filter (EKF). This filter is based upon the principle of linearization of the state transition matrix and the observation matrix with Taylor series expansions. Exploiting the assumption that all transformations are quasi-linear, the EKF simply makes linear all nonlinear transformations and substitutes Jacobian matrices for the linear transformations in the KF equations. The linearization can lead to poor performance and divergence of the filter for highly non-linear problems. Unfortunately the EKF has two important set backs. first, the in the derivation of the jacobian matrices the linear approximations can be complex. secondly this linearization can lead to filter instability in the time step intervals are not sufficiently small.

An improvement to the extended Kalman filter is the unscented Kalman filter (UKF). The basic premise behind the unscented kalman filters is that it is easier to approximate a Gaussian distribution than to approximate an arbitrary nonlinear function. it approximates the probability density resulting from the nonlinear transformation of a random variable. It is done by evaluating the nonlinear function with a minimal set of carefully chosen sample points. The posterior mean and covariance estimated from the sample points are accurate to the second order for any nonlinearity.

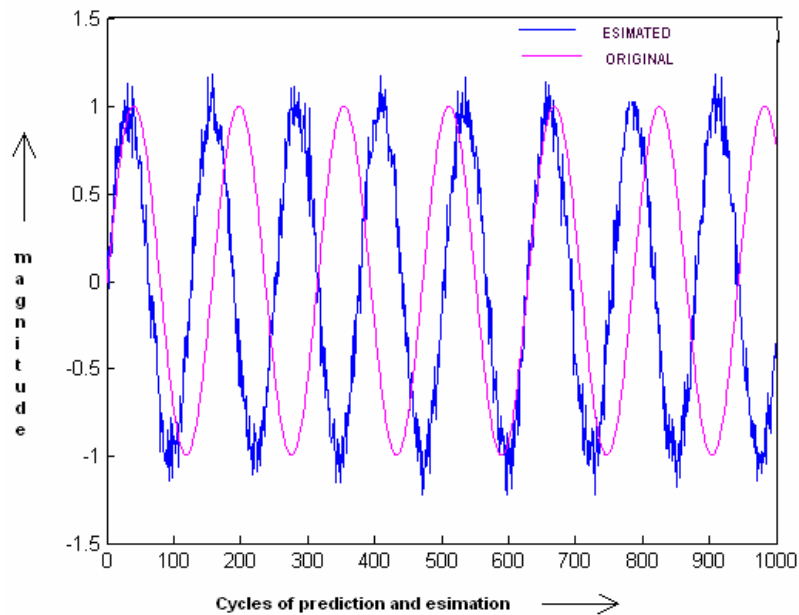
COMPARISON OF EXTENDED KALMAN AND UNSCENTED KALMAN FILTER



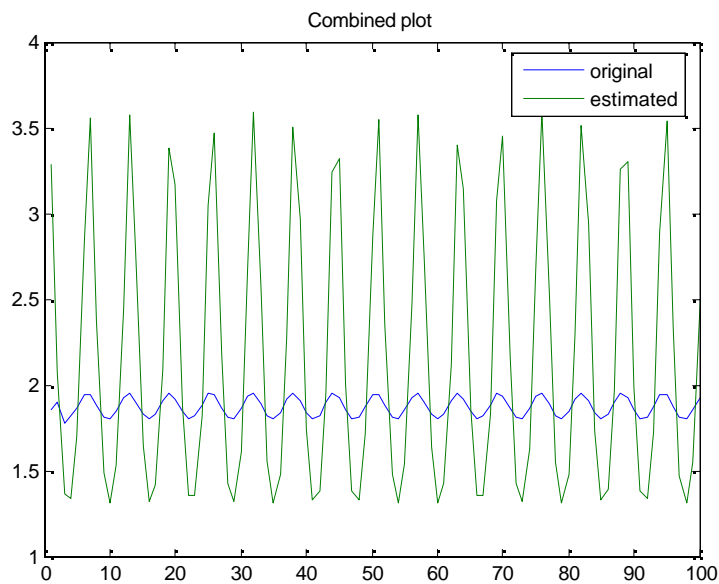
EXTENDED KALMAN FILTER AND UNSCENTED KALMANN FILTER ESTIMATE SIMULATION

ESTIMATE COMPARISON BETWEEN EXTENDED KALMAN AND UNSCENTED KALMAN FILTER (FOR A COMMON SIGNAL)

EXTENDED KALMAN FILTER PREDICTION :



UNSCENTED KALMAN FILTER ESTIMATE :



**ACCURATE PHASE FOLLOWING OF THE UNSCENTED KALMAN FILTER
WITH SOME AMPLITUDE ERROR.**

**EKF AND UKF FILTER ESTIMATE SIMULATION FOR COMMON
SIGNAL**

CONCLUSION

we have presented various alternatives to design an OPTIMAL FILTER.. We have studied the characteristics and uses of the KALMAN,EXTENDED KALMAN,UNSECNTED KALMAN AND THE EXTENDED COMPLEX KALMAN FILTERS.

The Kalman filter (KF) is an optimal linear estimator when the process noise and the measurement noise can be modeled by white Gaussian noise. In situations when the problems are nonlinear or the noise that distorts the signals is non-Gaussian, the Kalman filters provide a solution that may be far from optimal.

Nonlinear problems can be solved with the extended Kalman filter (EKF). This filter is based upon the principle of linearization of the state transition matrix and the observation matrix with Taylor series expansions. EKF has two important set backs. first, the in the derivation of the jacobian matrices the linear approximations can be complex. secondly this linearization can lead to filter instability in the time step intervals are not sufficiently small.

An improvement to the extended Kalman filter is the unscented Kalman filter (UKF). The basic premise behind the unscented kalman filters is that it is easier to approximate a Gaussian distribution than to approximate an arbitrary nonlinear function. The UKF consistently achieves a better level of accuracy than the EKF at a comparable level of complexity

This makes ECKF suitable for processing general complex-valued nonlinear and non stationary signals and bivariate signals with strong component correlations.

. We have demonstrated performance gain in a number of application domains, including state-estimation, dual estimation, and parameter estimation. Future work includes additional characterization of performance benefits, extensions to batch learning and non-MSE cost functions, as well as application to other architectures. However we only could analyze and study the filter's characteristics and their feasibility of implementation and not implement the filter to estimate the real time frequency due to time required in building the mat lab code to realize the actual filter.

REFERENCES

1. Simon J. Julier and Jeffrey K. Uhlmann. "Unscented Filtering and Nonlinear Estimation," *Proc. of the IEEE*, vol. 92, pp. 401-422, March 2004.
2. B. La Scala and R. Bitmead, "Design of an extended Kalman filter frequency tracker," *IEEE Trans. Signal Processing*, vol. 44, pp. 739-742, 1996
3. Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems," *Transaction of the ASME—Journal of Basic Engineering*, pp. 35-45, March 1960.
4. S. J. Julier and J. K. Uhlmann. "A new extension of the Kalman filter to nonlinear systems," *Proc. AeroSense: 11th Int. Symp. Aerospace/Defense Sensing, Simulation and Controls*, 1997, pp. 182-193.
5. B. La Scala, R. Bitmead, and B. G. Quinn, "An extended Kalman filter frequency tracker for high-noise environments," *IEEE Trans. Signal Processing*, vol. 44, pp. 431-434, 1996.
6. B. La Scala, R. Bitmead, and M. R. James, "Conditions for stability of the extended Kalman filter and their application to the frequency tracking problem," *Math. Control, Signals Syst.*, vol. 8, pp. 1-26, 1995.
7. A. A. Girgis and T. L. D. Hwang, "Optimal estimation of voltage phasors and frequency deviation using linear and nonlinear Kalman filtering," *IEEE Trans. Power App. Syst.*, vol. PAS-103, no. 10, pp. 2943-2949, 1984.
8. P. K. Dash, A. K. Pradhan, and G. Panda, "Frequency estimation of distorted power system signals using extended complex Kalman filter," *IEEE Trans. Power Del.*, vol. 14, no. 3, pp. 761-766, Jul. 1999.
9. Wan, E. A., and R. van der Merwe. The Unscented Kalman Filter for Nonlinear Estimation. In *Proceedings of Symposium 2000 on Adaptive Systems for Signal Processing, Communication and Control (AS-SPCC)*, IEEE Press, 2000.
10. Wan, E. A., and R. van der Merwe. The Unscented Kalman Filter, In *Kalman Filtering and Neural Networks*, S. Haykin (ed.), Wiley Publishing, 2001.